

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Etude de l'impact de l'information lors de la propagation d'une maladie dans une population

Rochez, Justine

Award date:
2011

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



**FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX
NAMUR**

Faculté des Sciences

Etude de l'impact de l'information lors de la propagation d'une maladie dans une
population

Mémoire présenté pour l'obtention
du grade académique de master en Sciences Mathématiques
Promoteur : Timoteo Carletti
Justine ROCHEZ
Juin 2011



**FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX
NAMUR**

Faculté des Sciences

Etude de l'impact de l'information lors de la propagation d'une maladie dans une
population

Mémoire présenté pour l'obtention
du grade académique de master en Sciences Mathématiques
Promoteur : Timoteo Carletti
Justine ROCHEZ
Juin 2011

Remerciements

Je tiens à remercier mon promoteur, Monsieur Timoteo Carletti, qui m'a permis d'étudier la diffusion dans les réseaux et sans qui ce mémoire n'aurait pas vu le jour.

Tout au long de cette année, ses conseils et sa disponibilité, malgré son emploi du temps chargé, m'ont permis d'avancer dans la direction, qui je l'espère est la meilleure.

Je le remercie encore pour avoir répondu patiemment à mes nombreuses questions et pour avoir participé en tant que lecteur de ce mémoire.

Mes remerciements vont aussi à Messieurs Jean-Charles Delvenne et Frank Callier, qui m'ont procuré une aide précieuse grâce à leur cours de Théorie des Graphes.

Je remercie également Madame Anne Lemaître et Messieurs Jean-Charles Delvenne et Renaud Lambiotte, lecteurs de ce mémoire.

Enfin, je voudrais remercier tout particulièrement ma famille et mon compagnon pour leur soutien et pour la patience dont ils ont dû faire preuve à certains moments.

Abstract

Alors que la théorie des graphes englobe des résultats fondamentaux sur les graphes, la théorie des réseaux s'intéresse aux graphes présents dans le monde réel, tel que les réseaux sociaux. Dans ces réseaux véhiculent modes, maladies, informations, etc.

L'objectif de ce mémoire est d'étudier la diffusion d'une maladie dans les réseaux sociaux lorsque l'information circule aussi. Nous avons choisi de travailler avec le modèle à agent qui consiste à effectuer une série de simulations sur un réseau modélisé au niveau microscopique. Grâce à ces simulations, nous avons pu étudier l'impact de l'information lors de la propagation de la maladie.

MOTS CLES : théorie des graphes - réseaux - diffusion - maladie - information

While the graph theory describes fundamental results on graphs, the network theory focuses on graphs of the real world, such as social networks. In this kind of network a lot of phenomena circulate, such as fashions, diseases, informations, etc.

The purpose of this thesis is to study the diffusion of diseases into social networks when the information is also diffusing. We choose to work with the agent based model which consists in simulating the diffusion of diseases in a modeled network at microscopic level.

These simulations will give us material in order to study the impact of the information in the presence of an epidemic.

KEY WORDS : graph theory - network - diffusion - disease - information

Table des matières

Introduction	2
1 Les modèles de diffusion	4
1.1 Diffusion de la maladie	4
1.1.1 Le modèle SI	5
1.1.2 Le modèle SIR	5
1.1.3 Le modèle SIS	6
1.2 Diffusion de l'information	6
1.2.1 Le modèle SI	6
1.2.2 Le modèle SIS	6
2 La modélisation d'un réseau	7
2.1 Définitions et concepts de base utiles à la modélisation d'un réseau	7
2.1.1 Réseau orienté et non-orienté	8
2.1.2 Matrice et liste d'adjacence d'un réseau	9
2.1.3 Degré d'un noeud	10
2.1.4 Réseau connexe	11
2.2 Modélisation d'un réseau	12
2.2.1 Réseau Random	13
2.2.2 Réseau Scale free	19
3 Modélisation de la propagation de la maladie et de l'information dans le réseau	25
3.1 Propagation de la maladie	25
3.1.1 Description	25
3.1.2 Pseudo-code	27
3.2 Propagation simultanée de la maladie et de l'information	29
3.2.1 Description	29
3.2.2 Pseudo-code	29
4 Analyse du modèle de propagation de la maladie et de l'information	31
4.1 Analyse du modèle de la propagation de la maladie sans information	31
4.1.1 Random	32
4.1.2 Scale-Free	39
4.2 Analyse du modèle de la propagation de l'information	45
4.3 Evolution du nombre de malades en l'absence d'information dans le réseau	47
4.3.1 Random	47

4.3.2	Scale-Free	54
4.4	Evolution du nombre de malades en présence d'information dans le réseau . . .	63
4.4.1	Random	63
4.4.2	Scale-Free	67
5	Comparaison de la propagation de la maladie avec et sans information	71
5.1	Introduction	71
5.2	Comparaison pour le réseau de type Random	72
5.3	Comparaison pour le réseau de type Scale-Free	75
	Conclusion	79
	Bibliographie	81
	Annexe	82

Introduction

Dans ce mémoire, nous étudierons la propagation de la maladie et de l'information dans une société virtuelle. En particulier, nous nous intéresserons à l'impact que peut avoir l'information lors de la propagation.

Ainsi, nous pourrions mesurer l'importance d'une campagne d'information lors d'une pandémie.

Pour étudier la propagation d'une maladie avec et sans information, nous avons plusieurs possibilités.

Il est possible d'étudier la propagation de la maladie et de l'information grâce aux équations différentielles en décrivant le taux de variation du nombre de personnes malades et celui des personnes saines en fonction de ces nombres et en faisant l'hypothèse de systèmes homogènes et bien dispersés.

Par contre, pour mettre mieux en évidence la structure de la société et le rôle que peut avoir chaque individu, nous avons choisi de travailler avec la méthode du modèle à agent où l'évolution de l'état de chaque individu est suivie dans le temps.

Dans la vie de tous les jours, nous trouvons des réseaux dans divers domaines. Nous pouvons citer, par exemples, les réseaux de cellules en biologie, les réseaux géophysiques, les réseaux écologiques, les réseaux internet, les réseaux financiers ou encore les réseaux sociaux.

Dans ce mémoire, nous allons nous intéresser aux réseaux sociaux.

Ces derniers représentent la structure dynamique d'une population. Dans les réseaux sociaux, plusieurs choses véhiculent. Les individus d'un réseau peuvent se passer une mode, des informations ou encore se transmettre des maladies via des virus et des bactéries.

Souvent lors de propagation de maladie, se diffuse en parallèle, l'information la concernant. Cela va permettre aux individus de mieux s'en protéger.

Nous modéliserons une société virtuelle dans laquelle nous ne ferons que diffuser la maladie dans un premier temps. Ensuite, en plus de la maladie, nous injecterons l'information pour constater l'impact que cette dernière peut avoir sur la propagation de la maladie.

Pour modéliser la structure de notre société, nous avons utilisé deux modèles différents. Le premier modèle est le modèle Random. Il correspond à la modélisation d'une population où chaque individu a en moyenne le même nombre de connections que les autres individus. Le second modèle correspond au modèle Scale-free. Cette modélisation correspond plus à la réalité. Elle représente une population où seuls quelques individus ont beaucoup de connections, les autres en ont beaucoup moins.

Pour effectuer la modélisation d'un réseau, nous avons utilisé le programme Matlab.

Le premier chapitre va introduire les différents modèles de diffusion d'une maladie et de l'information.

Il en existe principalement trois pour la maladie et deux pour l'information. Dans ce mémoire, nous avons choisi d'en utiliser un pour notre modélisation qui est le même pour la propagation de la maladie et pour celle de l'information.

Dans le second chapitre, nous expliquerons la modélisation de nos réseaux suivant le modèle Random et Scale-free.

Nous donnerons les définitions et concepts de base de la théorie des graphes utiles à la modélisation d'un réseau. Une fois la modélisation effectuée, nous donnerons une représentation des réseaux pour les deux modèles possibles.

Le troisième chapitre traitera de la propagation de la maladie seule et en présence d'information. Nous expliquerons comment nous procédons pour propager d'une part la maladie seule et d'autre part la maladie et l'information dans le réseau.

Dans le quatrième chapitre, nous effectuerons l'étude l'évolution du nombre de malades dans le réseau lorsque seule la maladie se propage et lorsque maladie et information se propagent simultanément.

Pour terminer, le cinquième chapitre relatera les différents résultats obtenus de l'impact de l'information sur la propagation de la maladie dans le réseau.

Chapitre 1

Les modèles de diffusion

Dans ce premier chapitre, nous expliquons les différents modèles de diffusion d'une maladie et de l'information¹.

1.1 Diffusion de la maladie

Toutes les maladies n'ont pas le même modèle de propagation. Certaines peuvent impliquer une infection permanente alors que d'autres aboutissent à une immunité totale des individus une fois rétablis. D'autres encore permettent une guérison temporaire.

Considérons un réseau défini par un ensemble de noeuds représentant les individus et un ensemble de liens représentant les interactions entre ces individus.

Dans ce réseau, se propage une maladie sur une période de temps.

A chaque temps t , l'état du système consiste en une partition de l'ensemble des noeuds parmi 3 sous-ensembles :

- $S(t)$: "susceptible nodes"
- $I(t)$: "infected nodes"
- $R(t)$: "removed nodes"

$S(t)$ correspond à l'ensemble des noeuds sains et susceptibles de tomber malade dans le réseau à l'instant t . $I(t)$ correspond à celui des noeuds infectés dans le réseau à l'instant t . Et enfin, $R(t)$ correspond à l'ensemble des noeuds guéris et immunisés dans le réseau à l'instant t . Ces 3 sous-ensembles vont évoluer différemment avec le temps, suivant le modèle de diffusion dans lequel on se trouve.

Il existe trois principaux modèles de diffusion différents

- le modèle SI
- le modèle SIR
- le modèle SIS

Nous les présenterons brièvement dans les sections suivantes.

1. La majorité des informations rassemblées dans ce chapitre sont issues du livre [2].

1.1.1 Le modèle SI

Le modèle SI correspond au modèle où une personne peut être soit saine (S), soit infectée (I).

Dans ce modèle, le sous-ensemble $R(t)$ est vide tout au long de la propagation de la maladie. A chaque temps t , chaque noeud sain du sous-ensemble $S(t)$ pourra devenir infecté et passer dans le sous-ensemble $I(t)$ si son voisinage l'est aussi.

Mais une fois infecté, le noeud ne pourra plus passer à l'état sain et pourra infecter son voisinage. L'état infecté est un état absorbant dans le modèle SI.

Ce modèle est assez flexible pour s'adapter à un large nombre de phénomènes. Celui qui nous intéresse est la contagion de certaines maladies dont la contamination se fait par le biais de contacts entre individus.

Pour ce modèle, nous pouvons reprendre l'exemple du Sida. Un fois infecté par la maladie, la personne ne pourra plus passer à l'état susceptible.

Le modèle SI s'avère aussi utile pour décrire la propagation des virus sur internet lors d'échanges de mails, ou encore la propagation de modes par imitation des voisins.

Dans ce mémoire, nous avons choisi de travailler avec ce type de modèle.

Nous l'avons choisi car il est à la base de ces trois modèles.

Une fois la modélisation effectuée pour ce type de modèle, des modifications sont possibles pour obtenir les deux autres.

1.1.2 Le modèle SIR

Le modèle SIR est basé sur le modèle SI, mais il envisage un état supplémentaire.

Cet état est l'état *recovery*, dont l'interprétation est simplement une situation dans laquelle le noeud en question ne peut plus être infecté ni transmettre l'infection.

Dans le cas de propagation de maladie, cet état peut représenter l'immunisé complète face à une maladie.

La Varicelle, par exemple, est une maladie qui se diffuse suivant le modèle SIR.

Durant son infection, un individu va pouvoir infecter son entourage. Mais une fois guéri, il sera immunisé et ne transmettra plus l'infection à son voisinage.

Cet état *recovery* peut aussi signifier "élimination", une élimination du noeud en raison de la mort de celui-ci, par exemple.

L'interprétation appropriée se fait suivant le contexte.

Si nous considérons que l'objet de diffusion est un certain type de comportement qui se propage par imitation, *recovery* peut signifier qu'un tel comportement est définitivement abandonné en raison de son inutilité et de sa dépréciation.

A la différence avec le modèle SI, l'infection n'est plus un état absorbant pour les noeuds. A chaque temps t , chaque noeud infecté du sous-ensemble $I(t)$ pourra guérir de manière irréversible et passer dans le sous-ensemble $R(t)$.

Quand cela arrive, le noeud en question ne pourra plus jamais transmettre l'infection à son voisinage.

1.1.3 Le modèle SIS

Le modèle SIS est semblable au modèle SI, mais à chaque temps t , chaque noeud infecté du sous-ensemble $I(t)$ pourra guérir de manière temporaire.

Une fois guéri, un noeud pourra de nouveau être exposé au risque de la contamination si son voisinage est infecté.

Dans un environnement biologique, cela se produit lorsque le rétablissement d'une maladie n'entraîne par une immunité permanente, c'est le cas pour la Grippe.

Une fois rétabli, un individu pourra de nouveau être contaminé et transmettre le virus. Dans un environnement social, une situation analogue se produit quand un certain comportement peut être abandonné, puis adopté de nouveau en fonction du comportement du voisinage.

1.2 Diffusion de l'information

Pour la diffusion de l'information, il existe deux modèles principaux de diffusion possibles. Le modèle SI et le modèle SIS.

Ce sont les même types de modèles que pour la diffusion de la maladie hormis le fait que l'objet de diffusion n'est plus la maladie mais l'information.

Nous ne parlerons plus de noeuds *infectés* mais de noeuds *informés*

Considérons un réseau dans lequel se propage, en plus de la maladie, l'information sur la même période de temps. A chaque instant t , l'ensemble des noeuds du système sera partitionné en deux sous-ensembles :

- $S(t)$: "susceptible nodes"
- $I(t)$: "informed nodes"

$S(t)$ correspond à l'ensemble des noeuds qui sont non-informés et susceptibles de devenir informés alors que $I(t)$ correspond à celui des noeuds qui sont informés.

Avec le temps, ces deux ensembles vont évoluer de manière différente selon le mode de propagation de l'information.

1.2.1 Le modèle SI

Le modèle SI correspond au modèle où un noeud du réseau est soit non-informés (S), soit informé (I).

A chaque temps t , un noeud non-informé pourra devenir informé si son voisinage l'est. Une fois informé, un noeud ne pourra plus passer à l'état susceptible.

L'état informé est un état absorbant dans le modèle SI.

Nous avons choisi de travailler avec ce type de modèle dans notre modélisation. Comme pour la maladie, il est le modèle de base pour la propagation de l'information.

1.2.2 Le modèle SIS

Le modèle SIS est semblable au modèle SI à la différence qu'une fois un noeud informé, il peut perdre cette information et être à nouveau non-informé et susceptible de devenir informé. L'état informé n'est plus un état absorbant pour le modèle SIS.

Chapitre 2

La modélisation d'un réseau

Dans ce chapitre, nous donnons une explication de la modélisation du réseau Random et du réseau Scale-free. Avant de passer à la modélisation, nous allons donner quelques définitions et concepts de base utiles à la modélisation d'un réseau. Ensuite, nous donnerons une explication de la modélisation d'un réseau de type Random et Scale-Free.

2.1 Définitions et concepts de base utiles à la modélisation d'un réseau

Dans cette section, nous allons définir quelques notions importantes¹ pour la modélisation d'un réseau.

Un réseau est défini par un ensemble de noeuds $N = \{1, 2, 3, \dots, n\}$ et par un ensemble de liens² $L \subseteq N \times N$. Un lien existant entre les noeuds i et j est noté (i, j) .

Le paramètre n , correspondant au nombre total de noeuds dans le réseau, est appelé *ordre* du réseau.

Graphiquement, un noeud est représenté par un point et un lien entre deux noeuds est représenté par un trait entre ceux-ci.

1. La majorité des informations rassemblées dans cette section sont issues du livre [1], ainsi que des cours de Théorie des graphes (cfr. [7] et [8]).

2. souvent appelées *arêtes*.

2.1.1 Réseau orienté et non-orienté

Une des caractéristiques d'un réseau est qu'il peut être orienté ou non.

Un réseau non-orienté est un réseau dont les liens sont dépourvus de sens. Dans ce cas, les liens (i, j) et (j, i) correspondent au même lien étant donné leur absence de sens sur les liens.

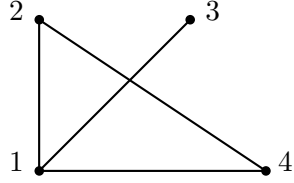


FIGURE 2.1 – Un réseau non-orienté.

La figure 2.1 correspond à la représentation graphique d'un réseau non-orienté d'ordre 4. Tous les liens sont dépourvus de sens. Le lien $(1, 2)$ correspond au même lien que $(2, 1)$.

Les liens d'un réseau orientés possèdent, quant à eux, un sens. Dans ce cas, le lien (i, j) correspond au lien dont l'origine est le noeud i et l'extrémité est le noeud j . Nous dirons alors que le noeud j est une connection du noeud i tandis que le noeud i est un parent de j .

De plus, l'existence du lien (i, j) ne va pas forcément impliquer l'existence du lien (j, i) .

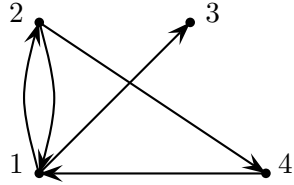


FIGURE 2.2 – Un réseau orienté.

La figure 2.2 représente un réseau orienté d'ordre 4. Tous les liens sont pourvus d'un sens. Nous constatons que le noeud 4 est une connection du noeud 2 alors que le noeud 1 est un parent du noeud 3.

Le lien $(1, 2)$ est distinct du lien $(2, 1)$.

2.1.2 Matrice et liste d'adjacence d'un réseau

La structure d'un réseau peut être caractérisée par sa matrice ou par sa liste d'adjacence. La matrice d'adjacence d'un réseau d'ordre n est définie par une matrice de taille $n \times n$, notée généralement A , où chaque élément a_{ij} est défini par

$$a_{ij} = \begin{cases} 1 & \text{si } (i, j) \in L \\ 0 & \text{sinon.} \end{cases}$$

A chaque réseau correspond, une matrice d'adjacence unique. Dans le cas de réseau non-orienté, la matrice d'adjacence sera une matrice symétrique.

Les matrices suivantes représentent les matrices d'adjacence des réseaux représentés par les figures 2.1 et 2.2 :

$$A_1 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

La matrice A_1 (respectivement A_2) correspond à la matrice d'adjacence du réseau représenté par la figure 2.1 (respectivement 2.2).

Nous pouvons constater que la matrice A_1 est symétrique, en effet, elle correspond à un réseau non-orienté.

La liste d'adjacence d'un réseau d'ordre n est définie par une liste de taille n , notée A_l , où chaque élément $a_l\{i\}$ est défini par :

$$a_l\{i\} = \{j \mid j \in N, (i, j) \in L\}$$

A chaque réseau correspond une liste d'adjacence unique.

Les listes suivantes correspondent aux listes d'adjacence des réseaux représentés par les figures 2.1 et 2.2 :

$$A_{l1} = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 4 \\ 1 \\ 1 & 2 \end{bmatrix} \quad A_{l2} = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 1 \end{bmatrix}$$

La liste A_{l1} (respectivement A_{l2}) correspond à la liste d'adjacence du réseau représenté par la figure 2.1 (respectivement 2.2).

Nous pouvons constater que dans la liste A_{l2} , l'élément $a_{l2}\{3\}$ est vide, en effet, dans la figure 2.2, le noeud 3 ne possède aucune connection.

2.1.3 Degré d'un noeud

La notion de degré d'un noeud est une notion très importante et très utile dans la modélisation d'un réseau.

Dans le cas des réseaux non-orientés, le degré d'un noeud, souvent noté k , correspond au nombre de liens qui lui sont incidents³.

Pour les réseaux orientés, la littérature différencie le degré entrant, noté k^{in} , et le degré sortant d'un noeud, noté k^{out} . Le degré entrant d'un noeud correspond au nombre de liens qui arrivent au noeud tandis que le degré sortant correspond au nombre de liens qui partent du noeud.

Dans la figure 2.1, le degré du noeud 1 est égal à 3 et celui du noeud 3 est égal à 1. En effet, il y a 3 liens incidents au noeud 1 alors qu'il n'y en a qu'un qui l'est au noeud 3.

La figure 2.2 représente un réseau dont le noeud 2 est de degré entrant égal à 1 alors que le degré sortant est égal à 2. Le noeud 3 a, quant à lui, un degré entrant égal à 1 et un degré sortant nul.

Il est possible de calculer le degré d'un noeud à partir de sa matrice ou liste d'adjacence. Si nous considérons un réseau non-orienté, le degré d'un noeud, à partir de la matrice d'adjacence du réseau, est la somme des éléments de la ligne (ou de la colonne) de la matrice correspondant au noeud. Le degré k_i du noeud i se calcule par

$$k_i = \sum_{j=1}^n a_{ij} = \sum_{j=1}^n a_{ji}$$

Par la liste d'adjacence, le degré d'un noeud correspond au nombre d'éléments qui composent la ligne de la liste d'adjacence correspondant au noeud, c'est-à-dire

$$k_i = \#(a_i\{i\})$$

Si nous considérons maintenant un réseau orienté, il va falloir différencier le degré entrant et sortant d'un noeud.

A partir de la matrice d'adjacence, le degré sortant d'un noeud est la somme des éléments de la ligne correspondant au noeud alors que le degré entrant d'un noeud est la somme des éléments de la colonne correspondante

$$k_i^{out} = \sum_{j=1}^n a_{ij} \ ; \ k_i^{in} = \sum_{j=1}^n a_{ji}$$

A partir de la liste d'adjacence, nous pouvons déduire facilement le degré sortant, ce qui n'est pas aussi évident pour le de gré entrant.

Le degré sortant d'un noeud correspond au nombre d'éléments qui composent la ligne correspondant au noeud,

$$k_i^{out} = \#(a_i\{i\})$$

Pour le degré entrant d'un noeud, il va falloir compter le nombre de lignes où le noeud apparaît.

$$k_i^{in} = \#\{j \mid j \in N, \ a_j\{i\} \ni i\}$$

3. Un lien est incident à un noeud si une de ses extrémités correspond à ce noeud.

2.1.4 Réseau connexe

Un réseau non-orienté est connexe s'il existe un chemin⁴ entre chaque couple de noeuds du réseau.

Dans le cas des réseaux orientés, un réseau sera connexe si le réseau non-orienté correspondant l'est.

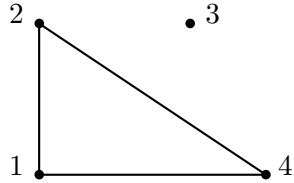


FIGURE 2.3 – Un réseau non connexe.

La figure 2.3 représente un réseau non connexe. Le noeud 3 n'est atteignable par aucun des noeuds du réseau.

La figure 2.1 représente, quant à elle, un réseau connexe. En effet, de tout couple de noeuds du réseau il est possible de trouver un chemin. Par exemple entre les noeud 2 et 3, un des chemins existant est 2, 1, 3. Il y existe d'autres, nous aurions pu citer, par exemple, le chemin 2, 4, 1, 3.

Pour vérifier qu'un réseau est connexe, nous allons utiliser son diamètre.

Dans ce mémoire, nous considérons que le diamètre D d'un réseau correspond à la distance⁵ maximale qu'il peut exister entre deux noeuds du réseau. Il existe d'autres définitions pour la notion de diamètre, comme celle de la distance moyenne entre tout couple de noeuds. Si le diamètre D est fini, alors le réseau sera connexe.

Le diamètre du réseau représenté par la figure 2.1 est 2, c'est la distance maximale qui existe entre deux noeuds du réseau.

Par contre le diamètre du réseau représenté par la figure 2.3 est infini. Il n'y a pas de chemin possible pour atteindre 3 quel que soit le noeud de départ.

4. Un chemin est une suite de noeuds telle que entre deux noeuds consécutifs du chemin, il existe un lien.

5. La distance entre deux noeuds correspond au nombre de liens qu'il faut emprunter au minimum pour aller d'un noeud à l'autre. S'il n'existe aucun chemin entre deux noeuds, la distance entre ces deux-ci est supposée infinie.

2.2 Modélisation d'un réseau

Dans cette section, nous allons passer à la modélisation d'un réseau. Notre modélisation de réseau sera une modélisation statique. Les relations entre les noeuds du réseau seront fixées et ne pourront plus être modifiées une fois déterminées. Dans la réalité, les connections d'un noeud sont en constante évolution. En effet, des amitiés se perdent et de nouvelles se créent sans cesse. Dans ce mémoire, nous avons fait le choix de simplifier la réalité en fixant les connections d'un noeud.

Nous avons vu dans la section précédente qu'un réseau pouvait être caractérisé par sa matrice ou par sa liste d'adjacence.

Dans ce mémoire, nous avons décidé de travailler avec les listes d'adjacence. Elles ont l'avantage que les connections de chaque noeud sont stockées directement contrairement aux matrices d'adjacence qui sont composées d'un grand nombre de 0. Tous ces 0 ont un coût en terme de stockage qu'il ne faut pas négliger surtout lorsque nous travaillerons avec des réseaux d'ordre élevé.

En Matlab, il est possible de créer des tableaux de cellules. Leur avantage est que chaque cellule pourra contenir un nombre différent d'objets, contrairement au tableau habituel où chaque ligne est de même longueur.

La commande permettant de créer un tableau de cellules est

```
noeud=cell(n,1)
```

noeud sera un tableau de cellules formé de n lignes et 1 colonne. Il correspond donc à un vecteur colonne où chaque cellule peut contenir des objets dont le nombre peut varier d'une cellule à une autre.

Il représente la liste d'adjacence du réseau où chaque cellule contiendra la liste des connections d'un noeud, $noeud\{i\}$ contenant la liste des connections du noeud i .

Une fois ce tableau de cellules créé, nous allons devoir déterminer les connections et donc le degré sortant de chaque noeud.

La détermination du degré des noeuds du réseau va dépendre du modèle choisi. Nous avons choisi d'en développer deux dans ce mémoire : le modèle Random et le modèle Scale Free.

Une fois le degré de chaque noeud fixé, nous pourrions attribuer une liste de connections à chaque noeud. Ces connections seront choisies aléatoirement parmi la liste de noeuds du réseau telles que le nombre de connections choisi pour un noeud correspondra à son degré et telles qu'un noeud ne puisse être connecté à lui même.

Etant donné que la détermination des noeuds se fait de manière aléatoire, la liste d'adjacence créée déterminera un réseau orienté⁶. Dans ce mémoire, le sens d'un lien entre deux noeuds correspondra au sens de transmission de la maladie et de l'information.

Nous avons supposé que la transmission de la maladie ne s'effectue pas de manière bilatérale et donc qu'un lien entre deux noeuds pourra être orienté.

Ceci est expliqué par le fait que lorsqu'un groupe de personnes se rencontrent, certaines personnes peuvent être malades et contaminer les autres. Etant donné que chaque personne a son

6. Il faudrait beaucoup de chance pour que la liste d'adjacence soit générée de manière à correspondre à un réseau non-orienté.

propre système immunitaire, chaque personne réagit différemment. Sans prendre de précautions particulières, certaines personnes sont moins sujettes à tomber malades.

Au sujet de la transmission de l'information, nous avons considéré le même schéma. Une personne, pouvant avoir des préjugés, ne tiendra pas compte de l'information que les autres lui donnent. La diffusion de l'information ne sera, elle non plus, pas symétrique.

Dans les sections suivantes, nous donnons une explication de la détermination des degrés ainsi que la détermination des listes de connections des noeuds en fonction de leur degré pour chacun des modèles (Random et Scale free).

2.2.1 Réseau Random

Un réseau Random est un réseau qui est généré par un processus aléatoire. Le premier modèle de réseau aléatoire a été popularisé par Paul Erdos et Alfred Renyi.

Dans ce mémoire, nous allons utiliser le processus binomial pour générer un réseau aléatoire. Le réseau Random est constitué à partir d'un nombre déterminé de noeuds et d'une probabilité qu'a un noeud de ce réseau d'être en lien avec un autre. L'idée de ce processus binomial est de considérer tous les couples de noeuds du réseau et de décider s'il y a ou non un lien entre le couple. Chaque lien entre un couple de noeuds du réseau aura une probabilité p d'exister, cela indépendamment du statut des autres liens.

Grâce à cette probabilité, nous allons déterminer le degré de chaque noeud.

Dans les sections suivantes, nous allons donner une description détaillée de la détermination du degré de chaque noeud pour le modèle Random. Nous donnerons également les détails sur la détermination des listes de connections des noeuds.

2.2.1.1 Description

Supposons que notre réseau soit composé de n noeuds.

Au maximum, un noeud ne peut avoir que $(n - 1)$ connections. En effet, nous avons supposé qu'un noeud ne peut être en lien avec lui-même.

Si nous comptabilisons le nombre maximum de liens qui peut exister dans un réseau de n noeuds, nous obtenons $\frac{n(n-1)}{2}$.

En effet, puisque chaque noeud a au maximum $(n - 1)$ liens et qu'il y a n noeuds dans le réseau, nous obtenons $n(n - 1)$ liens possibles.

Dans ce cas, un même lien entre deux noeuds est compté deux fois, une fois pour le premier noeud et une autre fois pour le noeud avec qui il est connecté. Nous devons donc diviser ce résultat par 2 et obtenons alors $\frac{n(n-1)}{2}$.

Si nous considérons maintenant qu'un noeud a une probabilité p d'être connecté à un autre, il possède alors en moyenne $p(n - 1)$ liens.

Pour l'ensemble du réseau, le nombre attendu de liens sera

$$E(m) = \frac{pn(n-1)}{2} \quad (2.1)$$

Dès lors, si nous savons que notre réseau est composé de n noeuds et m liens, il nous est possible de calculer la probabilité qu'a un noeud de notre réseau d'être connecté. Comme nous connaissons le nombre de liens qu'il y a dans le réseau, nous avons, grâce à l'équation 2.1,

$$m = \frac{p n (n - 1)}{2} \quad (2.2)$$

Il est alors évident que la probabilité qu'a un noeud d'être connecté se calcule par

$$p = \frac{2m}{n(n - 1)} \quad (2.3)$$

A partir de cette probabilité, il nous est possible de déterminer le degré de chaque noeud. Nous allons distribuer les degrés selon une loi de probabilité *Binomiale*($n - 1, p$).

La loi *Binomiale*(N, P) compte le nombre de succès lors de la répétition de N expériences, avec une probabilité P de réussite de l'expérience.

Dans notre cas, l'expérience sera de voir si un noeud en particulier est connecté avec un autre et un succès sera que les deux noeuds soient connectés.

Nous allons répéter l'expérience $(n - 1)$ fois, étant donné qu'un noeud peut être connecté avec au maximum $(n - 1)$ noeuds.

La probabilité de réussite est la probabilité p étant donné que cette probabilité p correspond à la probabilité d'être connecté à un noeud du réseau.

Une fois le degré sortant de chaque noeud choisi, nous déterminons les connections de ce noeud. Ce choix se fait aléatoirement parmi les noeuds du réseau.

Prenons pour exemple le noeud i qui a pour degré k . Nous choisirons alors aléatoirement k noeuds parmi les $(n - 1)$ restant pour constituer la liste de connections du noeud i .

Nous avons signalé précédemment qu'un réseau Random représente un réseau dont chaque noeud possède en moyenne le même nombre de connections que les autres noeuds du réseau. Théoriquement, il nous est possible de calculer la valeur du degré moyen ⁷ $\langle k \rangle$:

$$\langle k \rangle = \frac{2m}{n} \quad (2.4)$$

Ceci est assez évident compte tenu du fait que s'il y a m liens dans le réseau, un lien reliant deux noeuds, la somme des degrés de tous les noeuds sera $2m$. Pour avoir le degré moyen d'un noeud, il suffit alors de diviser ce résultat par n et nous obtenons le degré moyen $\langle k \rangle$.

Par l'équation 2.2, nous pouvons aussi exprimer le degré moyen $\langle k \rangle$ de la façon suivante

$$\langle k \rangle = \frac{2m}{n} = \frac{2(p \frac{n(n-1)}{2})}{n} = p(n - 1) \simeq pn$$

7. $\langle k \rangle$ va correspondre au degré moyen sortant étant donné que nous nous basons sur le nombre de connections que possède un noeud.

2.2.1.2 Pseudo-code

Dans cette section, nous allons donner les codes pour la modélisation d'un réseau de modèle Random.

Une fois le tableau de cellules *noeud* créé par la commande *noeud = cell(n, 1)*, nous allons déterminer le degré sortant de chaque noeud.

Pour déterminer la probabilité *p*, nous disposons, en paramètres, du nombre de noeuds *n* et du nombre de liens *m* du réseau à modéliser.

Voici le code pour déterminer le degré de chaque noeud

```
p=2*m/(n*(n-1));  
l=binornd(n-1,p,[1 n]);
```

Le calcul de la probabilité *p* s'est fait à partir de l'équation 2.3.

Le vecteur *l* est un vecteur ligne à *n* dimensions, *l(i)* correspondant au degré du noeud *i*.

Les degrés de chaque noeud sont attribués en fonction de la probabilité qu'a un noeud d'être connecté. Pour ce faire, nous avons utilisé la fonction *binornd(N,P,v)* de Matlab qui génère un vecteur de dimension *v* suivant la loi binomiale de paramètres *N* et *P*.

Le vecteur *l* est composé de nombres compris entre 0 et *n - 1*.

Une fois le degré de chaque noeud déterminé, nous pouvons spécifier la liste des connexions de chaque noeud.

Pour déterminer ces listes, nous allons procéder de la sorte :

```
for i=1:n  
if l(i)>0  
while length(noeud{i})<l(i)  
h=floor(rand()*n+1);  
if h~=i  
noeud{i}=union(noeud{i},h);  
end  
end  
end  
end
```

Nous allons parcourir tous les noeuds un à un.

Pour chaque noeud, si son degré est supérieur à 0, c'est-à-dire si le noeud a au moins une connection, et tant que nous n'avons pas atteint le degré de ce noeud, nous choisirons un noeud *h* qui pourra lui être connecté .

Ce noeud *h* est choisi aléatoirement parmi les nombres entre 1 et *n*.

Nous avons utilisé la commande Matlab *rand()* qui génère un nombre aléatoire situé dans l'intervalle (0, 1), dont la distribution est uniforme .

Pour que le nombre soit choisi entre 1 et *n*, nous avons multiplié *rand()* par *n* et l'avons incrémenté de 1. De cette manière, le nombre se situera dans l'intervalle $[1, n + 1[$.

Enfin, en ne prenant que la partie entière de ce nombre, nous nous assurons d'obtenir un entier

entre 1 et n .

Un noeud ne pouvant être connecté à lui même, nous ne rajouterons le noeud h à la liste de connections de i seulement si h est différent de i .

Nous avons utilisé la commande Matlab `union(A,B)` afin d'ajouter un noeud à une liste de connections. Cette commande permet d'ajouter l'ensemble B à l'ensemble A .

Une fois cette procédure effectuée pour tous les noeuds, nous sommes en possession d'un réseau de n noeuds et m liens qui est caractérisé par la liste d'adjacence *noeud*.

2.2.1.3 Représentation

Avant tout, donnons la distribution des noeuds par degré.
Considérons un réseau de 1000 noeuds et 10000 liens, nous obtenons

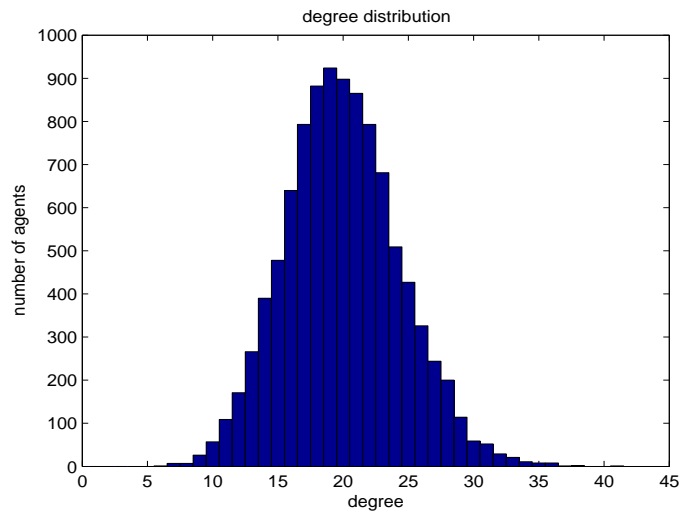


FIGURE 2.4 – Histogramme du nombre de noeuds par degré pour un réseau de 10000 noeuds et 100000 liens pour un réseau Random.

La figure 2.4 représente la distribution des degrés pour un réseau de type Random. Etant donné que la distribution des degrés a été faite suivant une loi binomiale, il n'est pas étonnant que cette distribution soit en forme de courbe de Gauss.

Traçons, en plus de la distribution des degrés, la distribution de la loi $Binomiale(n-1, p)$.

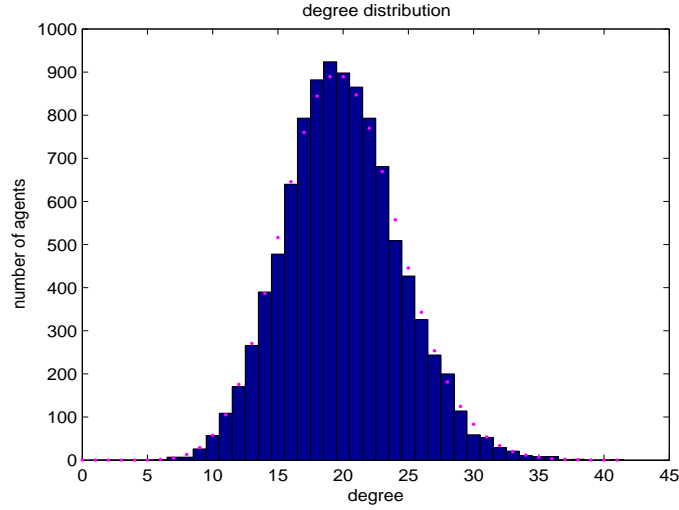


FIGURE 2.5 – Histogramme du nombre de noeuds par degré pour un réseau de 10000 noeuds à 100000 liens pour un réseau Random.

Sur la figure 2.5, la courbe tracée en rose correspond à la distribution d’une loi $Binomiale(n-1, p)$. La distribution des degrés des noeuds du réseau suit parfaitement la distribution théorique. Ceci n’est pas étonnant puisque les degrés ont été déterminés suivant cette loi.

Nous constatons que, sur les figures 2.4 et 2.5, la majorité des noeuds a un degré proche de 20.

Cette valeur correspond en réalité au degré moyen $\langle k \rangle$ du réseau.

En effet, étant donné que notre réseau est composé de 10000 noeuds à 100000 liens, de l’équation 2.4 , nous trouvons la valeur théorique du degré moyen

$$\langle k \rangle = \frac{2 \cdot 100000}{10000} = 20$$

Nous pouvons aussi représenter graphiquement le réseau qui a été défini grâce à la liste d'adjacence.

Pour représenter un réseau, nous allons utiliser le programme *Himmeli*. Il est disponible à partir d'internet⁸. Ce programme requiert plusieurs fichiers d'informations pour représenter le réseau, dont un sur la configuration du réseau, un autre sur les noeuds et un dernier sur les liens du réseau. Une fois en possession de ces fichiers, il est capable de générer une représentation du réseau.

Nous avons utilisé un programme Matlab⁹ qui, à partir de la matrice d'adjacence d'un réseau, crée les fichiers utiles au programme Himmeli.

Pour une question d'esthétique, nous avons choisi de représenter un réseau plus petit. La représentation d'un réseau à 10000 noeuds et 100000 liens se traduit par une immense tache tellement le nombre des noeuds et celui des liens sont grands. Nous avons donc réduit le nombre de noeuds et de liens pour obtenir quelque chose de visible.

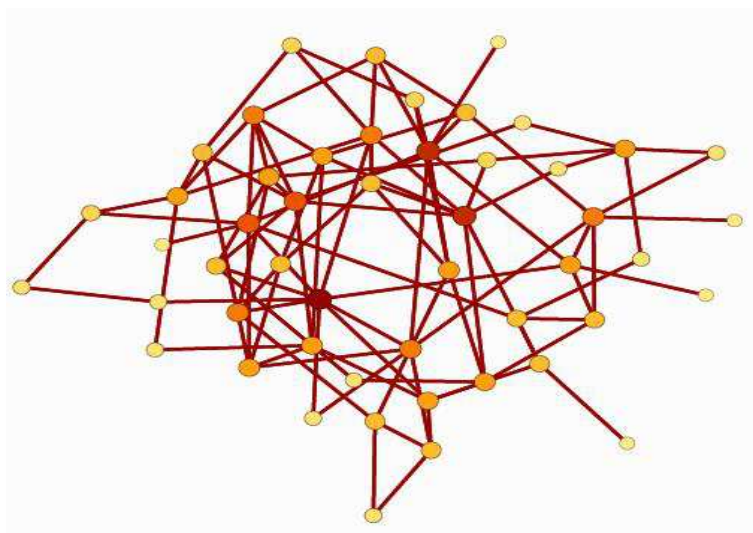


FIGURE 2.6 – Représentation d'un réseau Random de 50 noeuds et 100 liens.

Nous constatons sur la figure 2.6 que l'ensemble des noeuds est représenté par la même gamme de couleur. La couleur du point va dépendre du degré du noeud qu'il représente. Comme nous l'avons déjà dit, le réseau Random correspond à un réseau où l'ensemble des noeuds a en moyenne le même degré. Il donc évident que ces noeuds sont représentés par des couleurs de même ton.

Dans le cas d'un réseau de 50 noeuds et 100 liens, le degré moyen est

$$\langle k \rangle = \frac{2 \cdot 100}{50} = 4$$

8. <http://www.finndiane.fi/software/himmeli/>

9. Ce programme a été écrit par Timoteo Carletti.

2.2.2 Réseau Scale free

Un réseau Scale Free est un réseau dont la distribution des degrés suit une loi de puissance

Contrairement au réseau Random où chaque noeud a en moyenne le même nombre de connexions, dans un réseau Scale Free, certains noeuds possèdent un degré qui dépasse largement la moyenne. Ce type de réseau est plus conforme à la réalité.

Ces réseaux sont à la base du phénomène *Six degree Distribution*¹⁰ qui prétend que la distance moyenne entre deux personnes quelconques dans le monde est égale à six.

Le réseau Scale Free sera constitué à partir d'un nombre déterminé de noeuds mais également d'un paramètre γ . Ce paramètre va être utilisé pour attribuer un degré à chaque noeud. La distribution des degrés se fait selon une loi de puissance. Cette loi est caractérisée par $P(k) \sim \frac{1}{k^\gamma}$; la probabilité d'un noeud d'avoir un degré k est $\sim \frac{1}{k^\gamma}$. Cette loi de probabilité est définie pour toute valeur de k comprise dans l'intervalle $[1, n - 1]$. En effet, la définition de degré impose que celui-ci soit positif et la définition de la loi de probabilité Scale Free impose que le k soit distinct de 0. Chaque noeud aura donc au minimum une connection. De plus, un noeud, ne pouvant être en lien avec lui-même, aura un degré sortant maximum de $(n - 1)$.

Plus précisément, la distribution des degrés est définie par $P(k) = \frac{C_\gamma}{k^\gamma}$ où C_γ définit un coefficient de normalisation dépendant de γ . Nous pouvons donner une expression à ce coefficient.

Nous savons grâce aux propriétés des probabilités que la somme des probabilités d'une distribution est égale à 1 :

$$\sum_{k=1}^{k_{max}} P(k) = \sum_{k=1}^{k_{max}} \frac{C_\gamma}{k^\gamma} = C_\gamma \sum_{k=1}^{k_{max}} \frac{1}{k^\gamma} = 1$$

Nous obtenons alors l'expression du coefficient C_γ

$$C_\gamma = \frac{1}{\sum_{k=1}^{k_{max}} \frac{1}{k^\gamma}} \quad (2.5)$$

Une fois ce coefficient C_γ défini, nous attribuerons un degré à chaque noeud. Une fois le degré de chaque noeud fixé, une liste de connexions sera affectée à chaque noeud. Ces connexions seront choisies aléatoirement parmi l'ensemble des noeuds du réseau de manière à ce que le nombre de connexions, par noeud, corresponde au degré de chaque noeud.

Comme pour le cas Random, étant donné que la détermination des connexions se fait aléatoirement, le réseau créé sera orienté.

Comme précisé dans le paragraphe des réseaux Random, l'utilisation de réseau orienté permet de modéliser le fait que deux personnes ne réagissent pas de la même manière face à la propagation de la maladie et de l'information.

10. Pour plus d'informations, consulter le livre [4].

2.2.2.1 Description

Supposons que notre réseau soit un réseau composé de n noeuds et caractérisé par le paramètre γ .

Considérons notre fonction de probabilité des degrés qui est définie par

$$P(k) = \begin{cases} \frac{c_\gamma}{k^\gamma} & \text{pour } 1 \leq k \leq n-1 \\ 0 & \text{sinon.} \end{cases} \quad (2.6)$$

Considérons la fonction de répartition des degrés définie par

$$F(k) = \int_{-\infty}^k P(x) dx \quad (2.7)$$

$$= \int_1^k P(x) dx \quad (2.8)$$

Le passage de l'équation 2.7 à 2.8 se fait en utilisant l'équation 2.6.

Par définition de la fonction de répartition, $\forall k \in [1, n-1]$, nous aurons $F(k) \in (0, n-1]$. En effet, $F(1) > 0$ et $F(n-1) = 1$.

Dès lors, chaque valeur r comprise dans l'intervalle $(0, 1)$ pourra s'écrire sous la forme

$$r = F(k_i) = \int_1^{k_i} P(x) dx \quad (2.9)$$

Cette valeur r correspondra alors à un degré k_i qui sera distribué selon la loi de puissance considérée.

Chaque degré étant un entier fini, l'équation 2.8 peut s'écrire sous sa forme discrète

$$F(k) = \sum_1^k P(x) \quad (2.10)$$

Considérons un nombre aléatoire r distribué uniformément dans l'intervalle $(0, 1)$. Nous aurons

$$F(k) \leq r < F(k+1) \quad (2.11)$$

Par l'équation 2.10, nous trouvons

$$\sum_1^k P(x) \leq r < \sum_1^{k+1} P(x) \quad (2.12)$$

Pour chaque valeur aléatoire r , il va être possible de déterminer une valeur pour k qui vérifie l'équation 2.12.

Pour que cette valeur de k soit comprise entre les valeurs 1 et $n-1$, il faudra incrémenter la valeur trouvée de 1. En effet, si dès le départ, nous avons $F(1) > r$, le degré k déterminé par l'équation 2.12 sera égal à 0, or le degré est défini par une valeur strictement positive.

La valeur du degré, même incrémentée de 1, ne sera jamais égal à n .

Sinon, cela signifierait que $r = 1$, or la valeur de r a été choisie dans l'intervalle $(0, 1)$.

Une fois le degré déterminé pour chaque noeud, il nous faudra générer une liste de connexions.

Pour déterminer la liste de connections d'un noeud, nous allons procéder de la même manière que pour le cas Random. Nous allons choisir aléatoirement des noeuds parmi tous ceux du réseau de façon à ce qu'un noeud ne puisse être connecté à lui-même et que le nombre de connections soit égal au degré du noeud.

2.2.2.2 Pseudo code

Dans cette section, nous allons donner les codes pour la modélisation d'un réseau Scale-free.

Une fois le tableau de cellules *noeud* créé par la commande *noeud = cell(n,1)*, nous allons déterminer un degré à chaque noeud avant de lui attribuer une liste de connections.

Nous disposons du nombre de noeuds n dans le réseau ainsi que du paramètre γ .

Commençons par donner le code de la détermination du coefficient de normalisation C_γ

```
C=0;
for k=1:n-1
    C=C+(1/k^\gamma);
end
C=1/C;
```

Le coefficient C correspond au coefficient C_γ .

Une fois ce coefficient déterminé, nous pouvons passer à la partie détermination du degré de chaque noeud.

Pour éviter de calculer les probabilités cumulées pour chaque noeud, nous les calculons dès le départ et les enregistrons dans un vecteur

```
somme=C;
F=somme;
for i=2:n-1
    somme=somme+(C/i^\gamma);
    F=[F somme];
end
```

Le vecteur F contient l'ensemble des probabilités cumulées, le $i^{ème}$ élément contenant la probabilité cumulée $F(k=i)$. F est de dimension $n-1$ étant donné que $k \in [1n-1]$.

Pour chaque noeud du réseau, nous générons un nombre aléatoire compris entre 0 et 1.

```
r=rand(n,1)
```

Le vecteur r sera de dimension n . Il contiendra l'ensemble des nombres aléatoires des noeuds du réseau, $r(i)$ correspondant au nombre aléatoire du noeud i .

Une fois un nombre aléatoire attribué à chaque noeud, il faudra trouver le k pour chaque noeud qui vérifira l'équation 2.12

```

for i=1:n
    v=find(F<r(i));
    l(i)=length(v)+1;
end

```

La commande $\text{find}(X < y)$ renvoie les indices des éléments du vecteur X qui vérifie la condition $< y$.

Pour un noeud i , le vecteur v contiendra les indices des éléments du vecteur de probabilités cumulées F qui seront inférieurs au nombre aléatoire $r(i)$. La longueur du vecteur v incrémenté de 1, nous donne alors la valeur du degré du noeud i .

Le vecteur l contiendra l'ensemble des degrés qui auront été générés de cette façon pour l'ensemble des noeuds du réseau, $l(i)$ correspondant au degré du noeud i .

Une fois les degrés déterminés, nous pouvons déterminer la liste des connections de chaque noeud.

Pour ce faire nous procéderons de la même manière que pour le modèle Random.

```

for i=1:n
    if l(i)>0
        while length(noeud{i})<l(i)
            h=floor(rand()*n+1);
            if h~=i
                noeud{i}=union(noeud{i},h);
            end
        end
    end
end
end

```

Nous parcourons les noeuds du réseaux un à un.

Pour chaque noeud, si le degré de ce noeud est supérieur à 0, et tant que nous n'avons pas atteint le degré de ce noeud, nous lui choisissons un noeud h avec qui il pourra lui être connecté. Comme précédemment, ce noeud h est choisi aléatoirement par la commande $\text{floor}(\text{rand} * n + 1)$. Elle assure d'obtenir un nombre entier compris entre 1 et n .

Un noeud ne pouvant être en lien avec lui-même, nous ne rajouterons le noeud h à la liste de connections de i seulement si h et i sont distincts.

Une fois toute cette procédure effectuée, nous obtenons un réseau qui est caractérisé par une liste d'adjacence que nous avons générée.

Ce réseau sera composé de n noeuds et caractérisé par le paramètre γ .

2.2.2.3 Représentation

Représentons l'histogramme de la distribution des degrés pour un réseau de type Scale-free. Pour une meilleure compréhension, nous utilisons l'échelle logarithmique pour représenter la distribution des noeuds suivant les degrés.

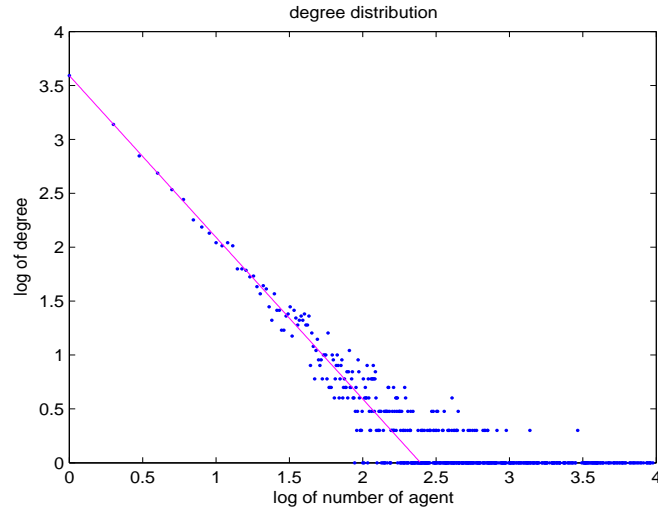


FIGURE 2.7 – Représentation logarithmique du nombre de noeuds par degré pour un réseau de 10000 noeuds avec un paramètre $\gamma = 1,5$.

La figure 2.7 montre bien que l'ensemble des noeuds ont un degré faible alors que seuls quelque uns ont un degré plus élevé que la moyenne. La droite représentée en rose sur la figure 2.7 représente la droite théorique selon laquelle les noeuds sont distribués. Nous constatons que la distribution des degrés suit cette droite.

Nous pouvons aussi donner une représentation graphique du réseau. Pour cela, nous utilisons toujours le programme *Himmeli*. Sur la représentation d'un réseau, la couleur des points qui représentent les noeud varient en fonction du degré du noeud.

Toujours pour une question d'esthétique, nous allons représenter un réseau de 50 noeuds dont le paramètre vaut 1,5.

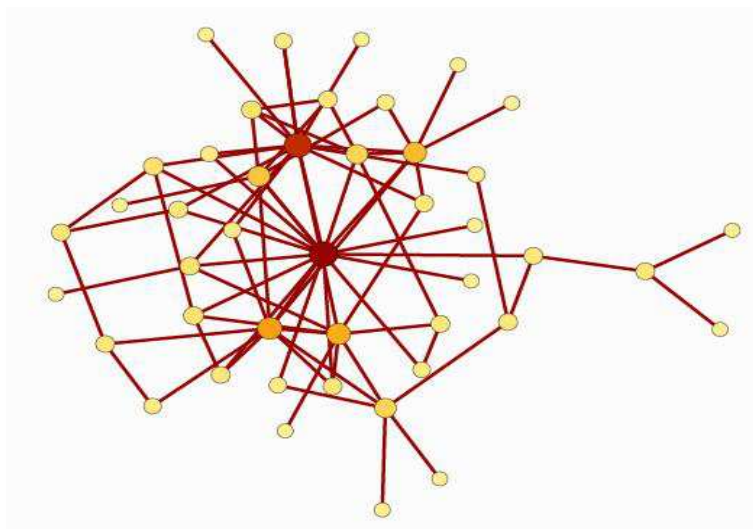


FIGURE 2.8 – Représentation d'un réseau Scale Free de 50 noeuds et avec $\gamma = 1.5$.

Nous remarquons, sur la figure 2.8, que seuls deux points sont plus gros que les autres. Ceci caractérise le fait que pour les réseaux Scale-free, très peu de noeuds possède un degré très grand.

Chapitre 3

Modélisation de la propagation de la maladie et de l'information dans le réseau

Dans ce chapitre, nous donnons une explication de la modélisation de la propagation de la maladie et de l'information dans notre réseau.

3.1 Propagation de la maladie

Dans cette section, nous donnerons une explication détaillée de la modélisation de la propagation d'une maladie dans un réseau. Cette modélisation est identique quelque soit le modèle de réseau, Random ou Scale Free. Le choix du modèle était important pour la détermination du degré sortant de chaque noeud.

3.1.1 Description

A ce stade, nous sommes en possession d'un réseau qui représente notre société virtuelle. A chaque noeud a été attribuée une liste de connections. Ces listes ont été générées de manière aléatoire suivant le degré sortant de chaque noeud.

Nous allons maintenant nous intéresser à la propagation de la maladie durant une période de temps.

Avant tout, un état de santé initial, sain ou malade, sera attribué à chaque noeud du réseau. Ces états de santé initiaux seront répartis de façon aléatoire en respectant une proportion de malades dans le réseau de départ.

Au fur et à mesure du temps, l'état de santé de chaque noeud va pouvoir évoluer ou non. Comme indiqué dans le chapitre 1, nous avons choisi de travailler avec un modèle de type SI. Un noeud sain pourra se retrouver infecté alors qu'un noeud malade restera malade.

Lorsque la répartition des états de santé de départ de chaque noeud est terminée, nous allons pouvoir procéder à la propagation de la maladie sur une certaine durée.

A chaque étape de temps, nous allons choisir un noeud aléatoirement dans le réseau. Supposons que le noeud i est choisi.

Si ce noeud i est sain, alors son état aura la possibilité d'évoluer. Etant donné que notre réseau est orienté et que la direction des liens indique le sens de propagation de la maladie, le noeud i ne pourra attraper la maladie que des noeuds qui ont une connection avec lui. Dans la suite de ce mémoire, nous appellerons ces noeuds les noeuds parents du noeud i .

Nous faisons ensuite l'hypothèse que si plus de la moitié des noeuds parents du noeud i sont malades, alors le noeud i tombera malade avec une certaine probabilité.

Par contre, si le noeud i est déjà malade, alors il le restera.

Pour étudier l'évolution des états de santé des noeuds, nous devons avoir accès à la liste des parents de chaque noeud. Pour ce faire, nous allons créer une deuxième liste, appelée liste "parentale", qui contiendra les listes des parents des noeuds du réseau. Cette liste parentale sera définie à partir de la liste d'adjacence d'un réseau.

Pour créer la liste des parents d'un noeud, il suffira de parcourir les listes de connections de tous les autres noeuds du réseau et de vérifier dans quelle liste figure ce noeud.

Le fait que le noeud j figure dans la liste du noeud i signifie que le noeud i est un parent de j . En parcourant l'ensemble des listes de connections, nous trouverons l'ensemble des parents d'un noeud.

La liste parentale d'un réseau d'ordre n , notée A_p , issue de la matrice d'adjacence A_l , est définie par une liste de taille n , où chaque élément $a_p\{i\}$ est défini par :

$$a_p\{i\} = \{j \mid j \in N, \quad a_l\{j\} \ni i\} \quad (3.1)$$

A chaque liste d'adjacence, et donc à chaque réseau, correspond une liste parentale unique. De plus, à partir de la liste parentale d'un réseau, le degré entrant d'un noeud se trouve immédiatement. En effet le degré entrant d'un noeud sera équivalent au nombre de parents que ce noeud possède

$$k_i^{in} = \#(a_p i)$$

Nous utiliserons cette notion de degré entrant, lorsque nous nous intéresserons à la propagation de la maladie parmi les noeuds ayant un certain degré entrant.

Reprenons, par exemple, la liste d'adjacence A_{l2} du réseau représenté par la figure 2.2.

$$A_{l2} = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 1 \end{bmatrix}$$

La liste parentale A_p correspondant à la liste d'adjacence A_{l2} sera

$$A_p = \begin{bmatrix} 2 & 4 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

En effet, en regard de la liste d'adjacence, les parents du noeud 1 sont les noeuds 2 et 4 puisque le noeud 1 figure dans la liste de connections de ces deux noeuds. Le noeud parent de 3 est 1 puisque le noeud 3 n'est une connection que du noeud 1. Nous pouvons aussi vérifier sur la figure 2.2 que cette liste parentale est bien celle du réseau qui y est représentée.

Une fois cette liste parentale créée, à chaque étape de temps, nous pourrions facilement voir si un noeud est infecté ou non, en fonction de l'état de ses parents.

3.1.2 Pseudo-code

Dans cette section, nous allons donner les codes pour la propagation de la maladie dans un réseau de n noeuds.

Une fois le tableau de cellules *noeud* complété, un réseau a été modélisé. Ce tableau de cellules *noeud* correspond à la matrice d'adjacence du réseau où chaque cellule représente une liste de connections. A partir de ce tableau de cellules *noeud*, nous allons en créer un autre qui correspondra à la liste parentale du réseau. Commençons par créer ce tableau de cellules

```
parent=cell(n,1)
```

Ce tableau de cellules *parents* sera de la même taille que *noeud*. Chaque noeud aura une liste de parents, *parent{i}* correspondant à celle du noeud i .

Une fois *parent* créé, il faudra le compléter. Pour ce faire, nous allons utiliser la définition de la matrice parentale exprimée à l'équation 3.1.

Le code Matlab correspondant est

```
for i=1:n
    for j=1:n
        if (length(find(noeud{j}==i))>0)
            parent{i}=union(parent{i},j);
        end
    end
    l2(i)=length(parent{i});
end
```

Nous allons parcourir les noeuds un à un afin de déterminer leur liste de parents. Supposons que l'on cherche la liste du noeud i . Nous allons parcourir la liste des connections de tous les noeuds du réseau. Nous rajouterons un noeud j comme parent du noeud i si ce noeud i fait partie de sa liste de connection.

La commande *find(noeud{j} == i)* va renvoyer l'ensemble des indices des éléments de *noeud{j}* qui seront égaux à i . Si la longueur de cet ensemble d'indices est supérieure à 0, c'est-à-dire si le noeud i figure dans la liste de connections du noeud j , alors nous rajouterons le noeud j à la liste des parents du noeud i grâce à la commande Matlab *union(A,B)* qui permet d'ajouter l'ensemble B à l'ensemble A déjà existant.

Une fois toutes les listes de connections pour un noeud vérifiées, nous pourrions enregistrer le degré entrant de ce noeud. Le vecteur $l2$ correspond au vecteur des degrés entrants de chaque noeud, $l2(i)$ correspondant au degré entrant du noeud i .

Une fois cette procédure effectuée pour l'ensemble des noeuds du réseau, nous serons en possession de la liste parentale du réseau ainsi que du vecteur de degré entrant de l'ensemble des noeuds du réseau.

Avant de procéder à la propagation de la maladie, nous devons attribuer un état de santé à chaque malade. Ces états de santé vont être déterminés aléatoirement suivant une proportion de malades *proportionmalade* de départ.

```
malade=binornd(1,proportionmalade,[1 n]);
```

Le vecteur *malade* contiendra l'ensemble des états de santé de chaque noeud, $malade(i)$ étant l'état de santé du noeud i .

Nous avons utilisé la commande Matlab *binornd*(N, P, v) qui génère un vecteur de dimension v , suivant la loi binomiale de paramètre N et P .

Ce vecteur *malade* sera constitué d'un ensemble de 0 et de 1, la valeur 1 correspondant à l'état de santé malade et la valeur 0 correspondant à l'état sain.

Une fois l'état de santé déterminé pour chaque noeud, nous allons procéder à la propagation de la maladie parmi les noeuds du réseau.

```
i=floor(rand*n + 1);
if (malade(i)==0)
    parentmalade=sum(malade(parent{i}));
    if parentmalade>l2(i)/2
        malade(i)=binornd(1,pmalade);
    end
end
```

A chaque étape de temps, nous allons sélectionner un noeud i parmi l'ensemble des noeuds du réseau grâce à la commande *floor*(*rand** $n + 1$) qui génère un nombre entier entre 1 et n . Si ce noeud i est sain ($malade(i) = 0$), il pourra devenir malade avec une probabilité *pmalade*. Nous comptabiliserons le nombre de parents malades du noeud i . Pour ce faire, nous considérons $malade(parent\{i\})$ qui ne donne que les états de santé des parents du noeud i . En effectuant la somme de tous les éléments de ce vecteur, nous aurons alors le nombre de malades parmi les parents du noeuds i . Si ce nombre *parentmalade* est supérieur à la moitié du nombre de ses parents, nous allons générer une valeur, 0 ou 1, telle que la probabilité que la valeur 1 apparaisse soit égale à *pmalade*.

Par contre si ce noeud i est déjà malade ($malade(i) = 1$) , alors il le reste.

3.2 Propagation simultanée de la maladie et de l'information

Dans cette section, nous donnerons une description détaillée de la propagation de la maladie lorsqu'en plus se propage l'information dans le réseau.

3.2.1 Description

Une fois en possession d'un réseau représentant notre société virtuelle, nous allons étudier la propagation de la maladie lorsque l'information la concernant se propage aussi dans le réseau. En plus d'attribuer un état de santé à chaque noeud, sain ou malade, nous lui attribuerons également un état d'information, informé ou non-informé.

Un noeud du réseau pourra donc être soit malade et informé, soit malade et non-informé, soit sain et informé et enfin sain et non-informé.

Comme pour l'état de santé, au fur et à mesure du temps, l'état d'information de chaque noeud pourra évoluer ou non. Etant donné que nous avons choisi de travailler avec le même modèle que pour la propagation de la maladie (le modèle SI), un noeud non-informé du réseau pourra le devenir tandis qu'un noeud déjà informé le restera.

Une fois les états de santé et d'information de chaque noeud répartis, nous allons procéder à la propagation de l'information et celle de la maladie sur une certaine durée.

A chaque étape de temps, nous allons sélectionner un noeud et voir si ses états de santé et d'information évoluent.

Nous allons procéder en deux étapes : une pour la propagation de l'information et une autre pour la propagation de la maladie.

Nous allons utiliser la même condition de propagation de l'information que celle utilisée pour la maladie : un noeud qui n'était pas informé pourra le devenir avec une certaine probabilité si plus de la moitié de ses parents le sont eux-mêmes.

La présence de l'information dans le réseau implique que si un noeud sain est informé, la probabilité qu'il devienne malade sera moins grande que s'il n'était pas informé.

Pour la propagation de l'information, nous devrons aussi avoir accès à liste parentale du réseau. Une fois en possession de cette liste, à chaque étape de temps, nous pourrions faire évoluer l'état d'information d'un noeud en fonction des états de ses parents.

3.2.2 Pseudo-code

Nous allons donner les codes pour la propagation de l'information ainsi que la propagation de la maladie lorsque l'information intervient.

A partir du réseau *noeud*, nous avons créé la liste parentale *parent* et nous avons déterminé le degré entrant de chaque noeud en fonction du nombre de leurs parents. Ces degrés entrants ont été enregistrés dans le vecteur *l2* (cfr section précédente).

L'élément $parent\{i\}$ contient la liste des parents du noeud i et l'élément $l2(i)$ correspond au degré entrant du noeud i .

En plus de l'attribution d'un état de santé à chaque noeud, nous leur attribuons un état d'information.

La détermination des états initiaux de santé et d'information se fait par

```
malade=binornd(1,proportionmalade,[1 n]);
info=binornd(1,proportioninfo,[1 n]);
```

Le vecteur *malade* sera composé des valeurs égales à 0, correspondant à l'état sain, et égales à 1, correspondant à l'état malade.

Le vecteur *info* sera lui aussi composé de valeurs égales à 0, correspondant à l'état non-informé, et égales à 1, correspond à l'état informé.

Ces deux vecteurs sont de dimension n , le $i^{ème}$ élément correspondant à l'état de santé ou l'état d'information du noeud i .

Nous allons ensuite passer à la propagation de l'information et de la maladie dans le réseau.

```
i=floor(rand*n+1);
if (info(i)==0)
    parentinfo=sum(info(parent{i}));
    if parentinfo>l2(i)/2
        info(i)=binornd(1,pinfo);
    end
end
if (malade(i)==0)
    parentmalade=sum(malade(parent{i}));
    if parentmalade>l2(i)/2
        if (info(i)==1)
            malade(i)=binornd(1,pmaladeinfo);
        else
            malade(i)=binornd(1,pmalade);
        end
    end
end
end
```

A chaque étape de temps, nous sélectionnons un noeud i du réseau aléatoirement.

Si ce noeud i n'est pas informé ($info(i) = 0$) et si la moitié de ses parents l'est ($parentinfo > l2(i)/2$), alors il pourra devenir informé avec une probabilité $pinfo$.

Ensuite, si le noeud est sain ($malade(i) = 0$), son état de santé pourra être modifié si plus de la moitié de ses parents est malade ($parentmalade > l2(i)/2$). La probabilité qu'il aura de tomber malade variera selon son état d'information. S'il est informé, la probabilité qu'il a de tomber malade, $pmaladeinfo$, sera moins grande que celle qu'il aura s'il n'est pas informé, $pmalade$.

Chapitre 4

Analyse du modèle de propagation de la maladie et de l'information

Dans ce chapitre, nous allons analyser la propagation de la maladie avec et sans information dans le réseau.

4.1 Analyse du modèle de la propagation de la maladie sans information

Lors d'une simulation de propagation de la maladie dans une population sans information, le caractère stochastique du modèle intervient lors de

- la création d'un réseau, soit Random, soit Scale Free
- la détermination des états de santé initiaux
- la propagation de la maladie

Afin de réduire cette part de stochastique, nous avons fait le choix de fixer un réseau de type Random *random1* et un réseau de type Scale-Free *scale1* que nous utiliserons pour toutes les simulations que nous effectuerons dans la suite de ce mémoire.

Ces deux réseaux ont été choisis selon certains critères. Le réseau *random1* a été modélisé suivant les paramètres $n = 1000$ et $m = 10000$. Le réseau *scale1*, quant à lui, a été modélisé selon les paramètres $n = 1000$ et $\gamma = 2$. De plus, nous avons vérifié que ces deux réseaux soient connexes. Rappelons que pour qu'un réseau soit connexe, son diamètre doit être fini. Pour chacun des réseaux, nous avons donc calculé l'ensemble des distances qui existent entre tout couple de noeuds du réseau. Le diamètre correspond à la distance maximale qui existe entre ces couples. La distance entre deux noeuds d'un couple est calculée grâce à l'algorithme de Dijkstra qui calcule le plus court chemin entre deux noeuds. Nous avons utilisé une fonction Matlab¹ qui calcule le diamètre d'un réseau, suivant cet algorithme, sur base de la liste d'adjacence d'un réseau.

1. Cette fonction *DiaL2m* a été écrite par Timoteo Carletti.

Une fois en possession des deux types de réseaux, *random1* et *scale1* connexes, nous allons effectuer une série de simulations à partir de ceux-ci. Nous allons étudier l'influence des distributions des malades de départ sur la dynamique de la maladie. Nous effectuerons cette étude pour chacun des deux réseaux que nous venons de fixer et en considérant les proportions des malades au départ 0.3, 0.4 et 0.5. Il est vrai que considérer une population dont 30% sont déjà malades au départ est énorme mais à partir d'une proportion inférieure et avec notre hypothèse de diffusion de la maladie², la maladie se propagerait très peu, voire pas du tout.

Nous allons maintenant présenter les différents résultats obtenus pour les réseaux *random1* et *scale1*.

4.1.1 Random

Pour cette partie, nous considérons le réseau *random1* qui a été modélisé. Nous allons effectuer une série de simulations pour les différentes proportions de malades au départ et en considérant que la probabilité de tomber malade est de 0.8.

Avec une proportion de malades de départ de 0.3

Tout d'abord, nous considérons que la population comporte 30% de malades au départ. Pour chacune des simulations que nous allons effectuer, nous considérerons une nouvelle distribution des malades. Après avoir effectué 15 simulations pour étudier l'évolution du nombre moyen de malades pour l'ensemble de ces simulations, nous obtenons

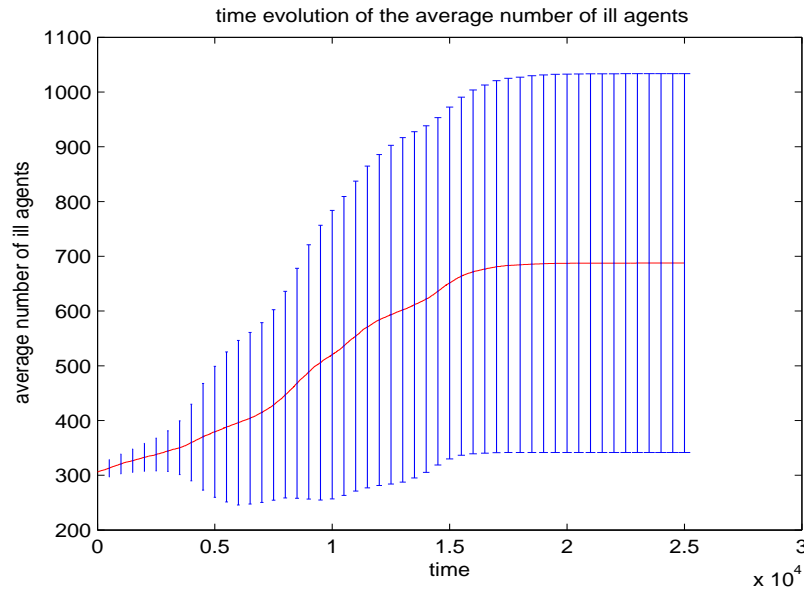


FIGURE 4.1 – Evolution du nombre moyen de malades pour le réseau *random1* avec 30% de malades au départ.

2. La contamination d'un noeud a lieu si plus de la moitié de ses noeuds parents sont atteints par la maladie.

La figure 4.1 représente l'évolution du nombre moyen de malades pour les 15 simulations effectuées. En plus de la courbe correspondant à l'évolution du nombre moyen de malades, représentée en rouge, nous avons représenté les déviations standards qui correspondent à la dispersion des mesures autour de la moyenne. Nous pouvons constater que les mesures se dispersent beaucoup autour de la courbe moyenne lorsque nous considérons une nouvelle distribution des malades à chaque simulation. La distribution des malades au départ influence donc beaucoup la dynamique de la maladie. D'une simulation à l'autre, les résultats de la propagation de la maladie sont très différents. Nous avons donc choisi de fixer des distributions des malades afin que les résultats, lors d'une série de simulations, soient similaires entre eux.

Lors des simulations effectuées, nous avons pu distinguer 2 cas distincts de propagation de la maladie selon les distributions de malades de départ. Le premier cas correspond à la situation où, selon une distribution des malades de départ, la propagation de la maladie est totale, c'est-à-dire que l'ensemble de la population est touchée par la maladie après un certain temps. La deuxième situation correspond à celle où seule une partie de la population est malade, même si la durée est longue. Dans ce cas, nous parlerons de propagation partielle de la maladie.

Nous avons isolé deux distributions des malades de départ dans le réseau qui aboutissent à ces deux situations différentes. La première distribution, *rmalade3_1*, correspond à la distribution qui va engendrer une propagation totale de la maladie. La deuxième distribution, *rmalade3_2*, correspond à celle qui engendre une propagation partielle de la maladie.

Considérons la distribution des malades au départ *rmalade3_1*. Si nous effectuons 15 simulations de la propagation de la maladie dans le réseau, où chaque simulation conserve cette distribution de malades, nous obtenons

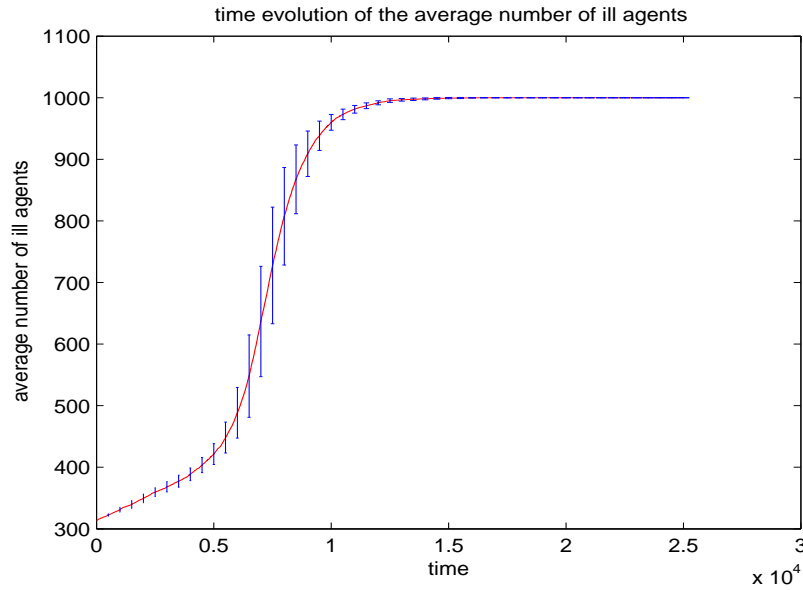


FIGURE 4.2 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution des malades *rmalade3_1*.

Nous constatons grâce à la figure 4.2 que cette distribution *rmalade3_1* engendre bien une propagation totale de la maladie pour le réseau *random1*. De plus, nous constatons que les résultats ne se dispersent plus beaucoup autour de la moyenne. Chaque simulation donne des résultats similaires.

Pour la distribution de malades de départ *rmalade3_2*, si nous effectuons toujours 15 simulations de propagation de la maladie, où chaque simulation conserve cette distribution de malades, nous obtenons

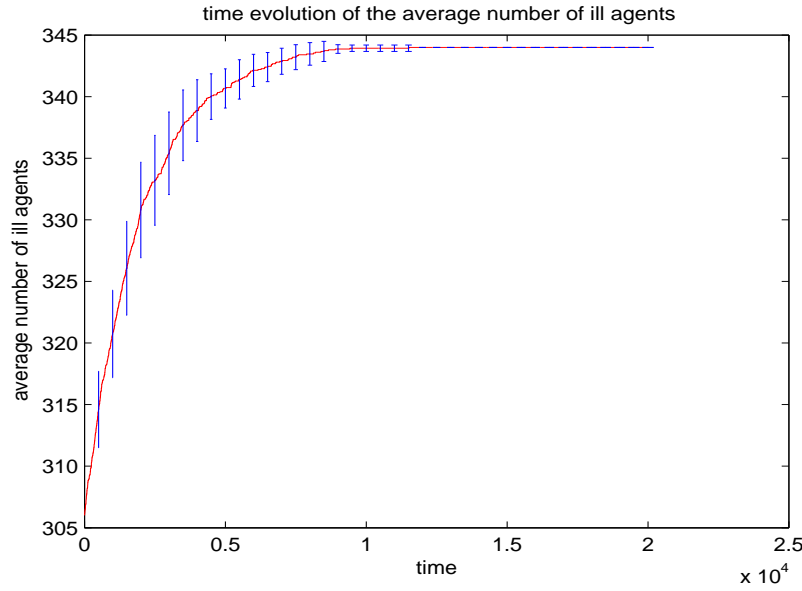


FIGURE 4.3 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution de malades *rmalade3_2*.

La figure 4.3 représente l'évolution de nombre moyen de malades lorsque la distribution *rmalade3_2* est utilisée. Nous constatons que cette distribution engendre bien une propagation partielle de la maladie. De plus, les résultats obtenus lors de chaque simulation se dispersent peu autour de la moyenne.

Nous pouvons déduire que, pour une proportion de malades au départ égale à 0.30, une fois la distribution des malades fixée, la dynamique de la maladie ne change plus. Dans la suite de ce mémoire, lorsque nous travaillerons avec le cas *random1* et une proportion de malades au départ de 0.30, nous considérerons ces deux distributions *rmalade3_1* et *rmalade3_2*.

Avec une proportion de malades de départ de 0.4

Nous considérons maintenant que la population comporte 40% de personnes malades au départ dans le réseau.

Nous allons effectuer une série de 15 simulations de la propagation de la maladie avec une distribution de malades différente pour chaque simulation.

Nous obtenons

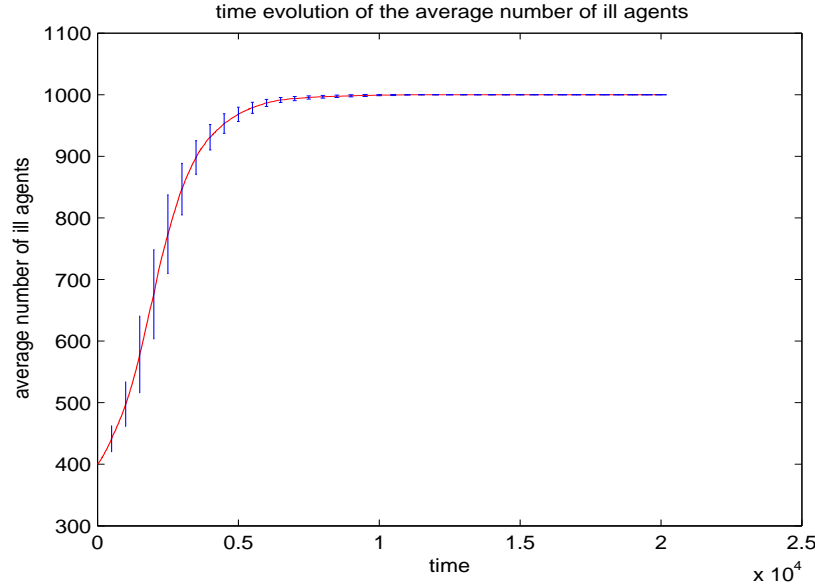


FIGURE 4.4 – Evolution du nombre moyen de malades moyen pour le réseau *random1* avec 40% de malades au départ.

La figure 4.4 représente l'évolution du nombre moyen de malades pour les 15 simulations de la propagation de la maladie effectuées avec une nouvelle distribution des malades à chacune des simulations. En plus de la courbe moyenne, nous avons représenté les déviations standards. Nous pouvons remarquer que lorsque 40% de la population sont malades au départ, les mesures se dispersent peu autour de la courbe moyenne et que la propagation de la maladie est totale.

La distribution des malades de départ dans le réseau, lorsque 40% de la population sont malades, n'a donc pas une grande influence sur la dynamique de la maladie. Quelque soit la distribution des malades avec une proportion de malades de 0.4, les résultats de chaque simulation sont similaires. Nous allons donc fixer une distribution de malades de départ que nous appellerons *rmalade4*

Nous effectuons 15 simulations de la propagation de la maladie, où chaque simulation conserve cette distribution de départ et nous obtenons

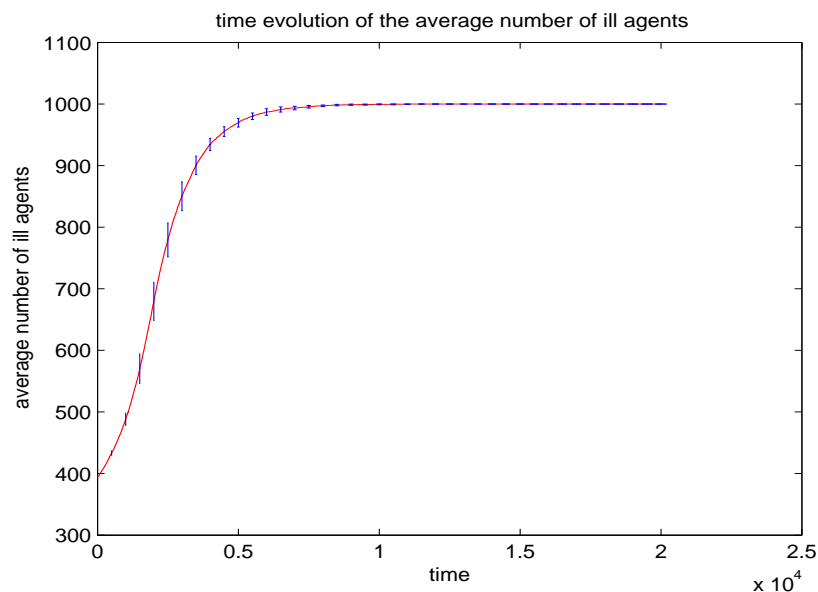


FIGURE 4.5 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade4*.

La figure 4.5 représente l'évolution du nombre moyen de malades lorsque la distribution *rmalade4* est utilisée à chaque simulation.

Cette distribution engendre bien une propagation totale de la maladie. Une fois la distribution fixée, les résultats de chaque simulation ne varient pratiquement plus.

Dans la suite, lorsque nous travaillerons avec le réseau *random1* et avec une proportion de 40% de personnes malades au départ, nous utiliserons cette distribution de malades *rmalade4*.

Avec une proportion de malades de départ de 0.5

Considérons maintenant que la population comporte une proportion de malades au départ de 0.5 et effectuons une série de 15 simulations de la propagation de la maladie, où chaque simulation utilise une nouvelle distribution de malades. Nous obtenons

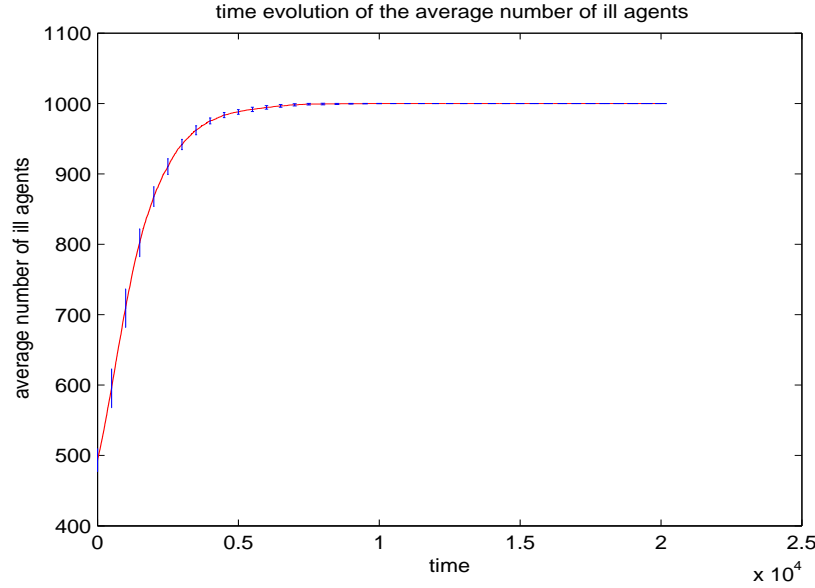


FIGURE 4.6 – Evolution du nombre moyen de malades pour le réseau *random1* avec 50% de malades au départ.

La figure 4.6 représente l'évolution du nombre moyen de malades lors des 15 simulations de la propagation de la maladie avec une nouvelle distribution des malades à chaque simulation. Nous constatons que, comme pour la proportion de 0.4, les mesures se dispersent peu autour de la courbe moyenne. L'influence de la distribution des malades n'est donc pas importante pour la dynamique de la propagation de la maladie. De plus, nous constatons que la propagation de la maladie est totale.

Etant donné que quelque soit la distribution considérée, les résultats de chaque simulation sont similaires, nous allons fixer une distribution de malades de départ que nous appellerons *rmalade5*.

Nous effectuons ensuite une série de 15 simulations de la propagation de la maladie sur base de cette seule distribution *rmalade5*.

Nous obtenons

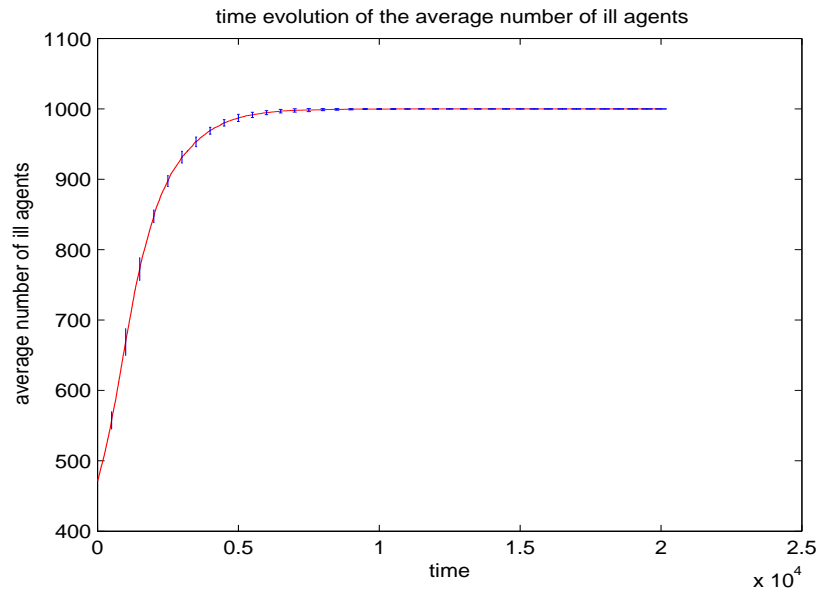


FIGURE 4.7 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade5*.

La figure 4.7 représente l'évolution du nombre moyen de malades lorsque la distribution *rmalade5* est utilisée pour chaque simulation.

Cette distribution engendre bien une propagation totale de la maladie. De plus, pour chacune des simulations, nous obtenons des résultats similaires. Une fois la distribution fixée, les résultats de chaque simulation ne varient pratiquement pas.

Dans la suite, lorsque nous considérerons une population modélisée par *random1* et dont 50% de la population est infectée au départ, nous utiliserons la distribution de malades *rmalade5*.

4.1.2 Scale-Free

Pour cette partie, nous considérons le réseau *scale1* qui a été modélisé. Comme pour le cas de Random, nous allons effectuer une série de simulations de propagation de maladie pour les différentes proportions de malades au départ que nous avons choisie et en considérant la probabilité de tomber malade égale à 0.8.

Avec une proportion de malades de départ de 0.3

Nous considérons que notre population, représentée par le réseau *scale1*, possède une proportion de malades au départ de 0.3.

Nous allons effectuer 15 simulations de propagation de la maladie, en considérant une nouvelle distribution de malades de départ à chaque simulation. Nous obtenons

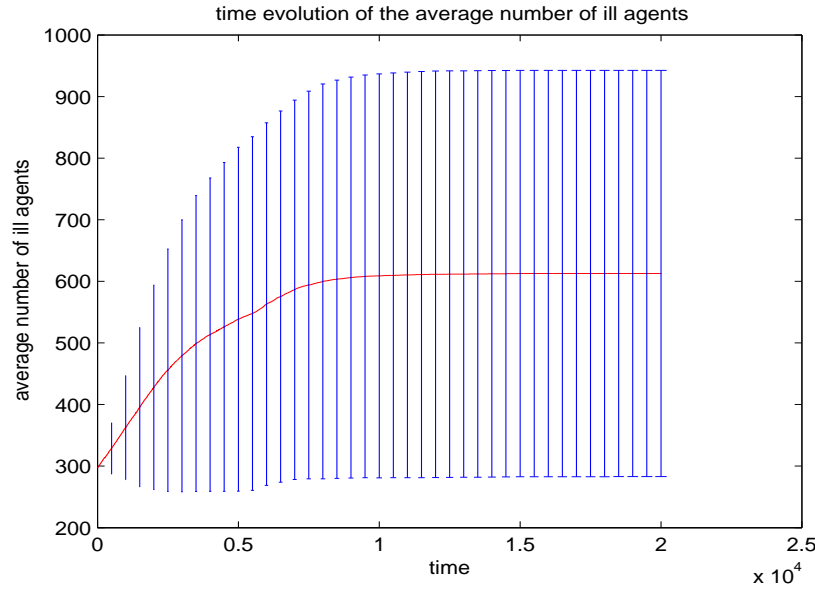


FIGURE 4.8 – Evolution du nombre moyen de malades pour le réseau *scale1* avec 30% de malades au départ.

La figure 4.8 représente l'évolution du nombre moyen de malades lors des 15 simulations de la propagation de la maladie, ainsi que la dispersion des résultats autour de cette courbe moyenne.

Nous pouvons constater que lorsqu'une nouvelle distribution des malades est considérée à chaque simulation, les mesures se dispersent beaucoup autour de la courbe moyenne. Nous pouvons en déduire que la distribution des malades influence de manière considérable la dynamique de la maladie. D'une simulation à l'autre, les résultats sont très différents. Nous avons donc fait le choix de fixer des distributions pour que les résultats lors d'une série de simulations soient similaires.

Lors de ces simulations, nous avons pu distinguer 2 cas différents de propagation de la maladie selon les distributions de malades. Certaines distributions de malades engendrent une propagation totale de la maladie parmi la population alors que d'autres engendrent une propagation partielle.

Nous avons isolé deux de ces distributions qui amènent à ces deux cas distincts. La première distribution *smaladie3_1* correspond à la distribution qui engendre l'infection de la totalité de la population alors que le distribtuion *smalade3_2* correspond à celle qui engendre l'infection partielle de la pouplation.

Considérons tout d'abord la distribution de malades *smalade3_1*. En utilisant cette distribution pour chacune des 15 simulations que nous effectuons, nous obtenons

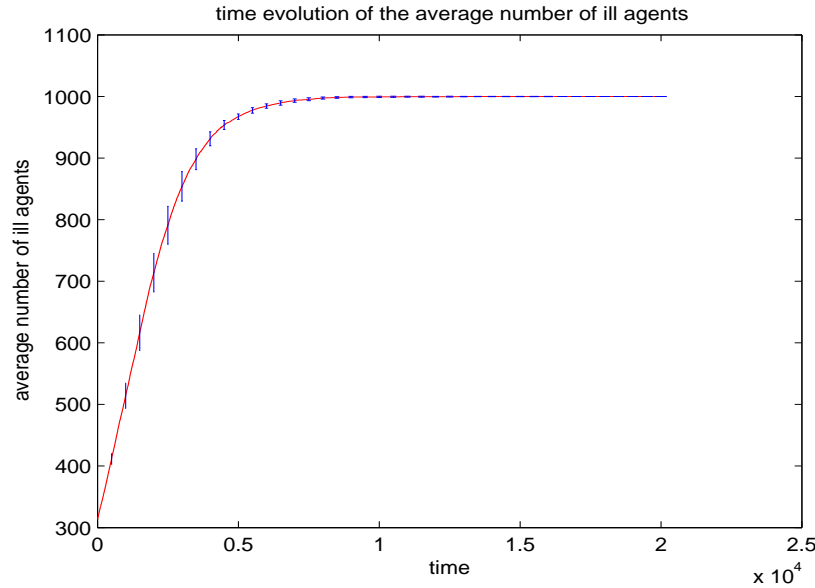


FIGURE 4.9 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade3_1*.

Nous constatons, grâce à la figure 4.9, que la distribution *smalade3_1* engendre bien une propagation totale de la maladie. De plus, les résultats des différents simulations ne se dispersent pratiquement pas autour de la courbe moyenne. Une fois cette distribution fixée, la dynamique de la maladie reste similaire pour chacune des simulations.

Ensuite, considérons la distribution des malades *smalade3_2*. Comme précédemment, nous effectuons 15 simulations de la propagation de la maladie en utilisant cette distribution à chaque simulation.

Nous obtenons

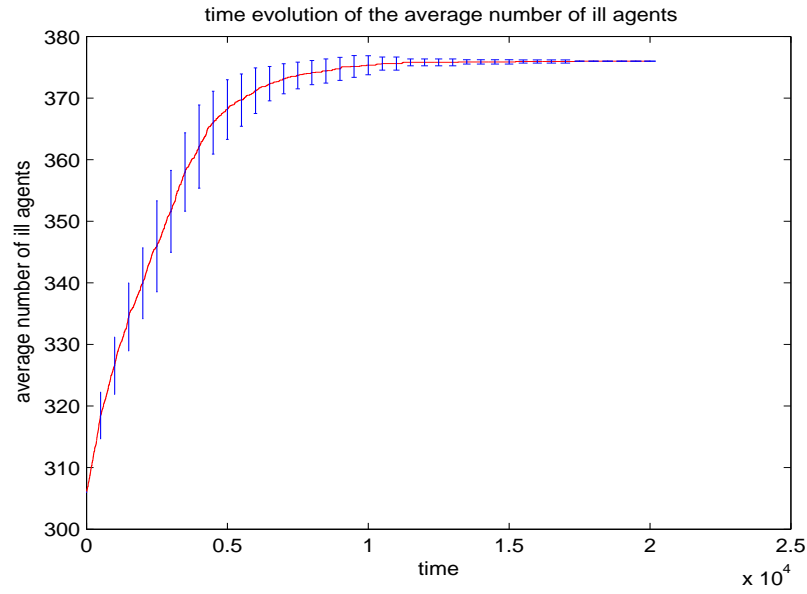


FIGURE 4.10 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade3_2*.

La figure 4.10 représente l'évolution du nombre moyen de malades en considérant la distribution *smalade3_2* à chaque simulation. Nous pouvons constater que cette distribution amène à une propagation partielle de la maladie dans la population. Les résultats des différentes simulations se dispersent un peu plus mais restent néanmoins proches de la courbe moyenne.

Avec une proportion de malades de départ de 0.4

Nous allons maintenant considérer des distributions de malades une proportion de malades de départ de 0.4.

Effectuons 15 simulations de la propagation de la maladie dans le réseau avec pour chacune des simulations une distribution différente des malades. Nous obtenons

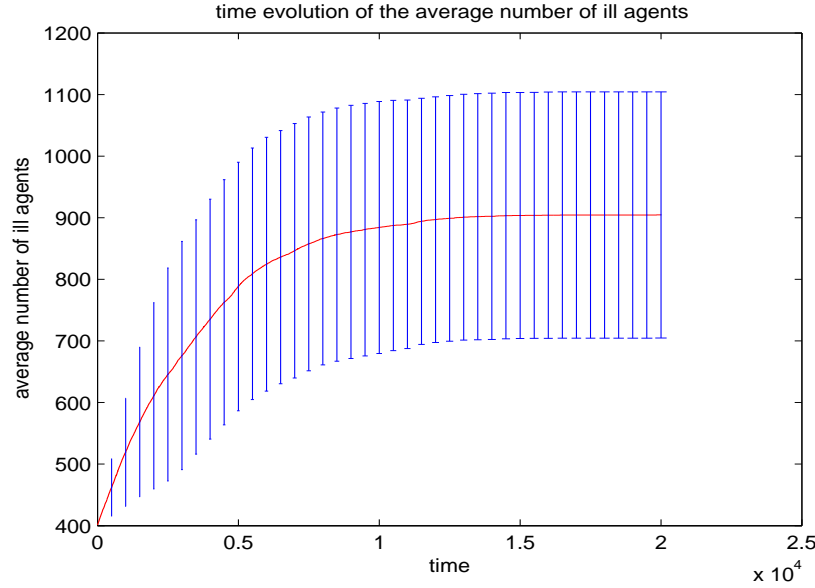


FIGURE 4.11 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade4*.

Grâce à la figure 4.11, nous pouvons constater que les résultats des différentes simulations se dispersent beaucoup autour de la courbe moyenne. La distribution des malades a donc une influence sur la dynamique de la maladie. D'une simulation à l'autre, les résultats sont différents.

Lors de ces simulations effectuées, nous avons pu isoler deux distributions qui amènent à des résultats distincts. La première *smalade4_1* amène à une propagation totale de la maladie alors que la deuxième *smalade4_2* amène à une propagation partielle.

Considérons ces deux distributions *malade4_1* et *malade4_2* et effectuons 15 simulations de la propagation de la maladie. Nous obtenons

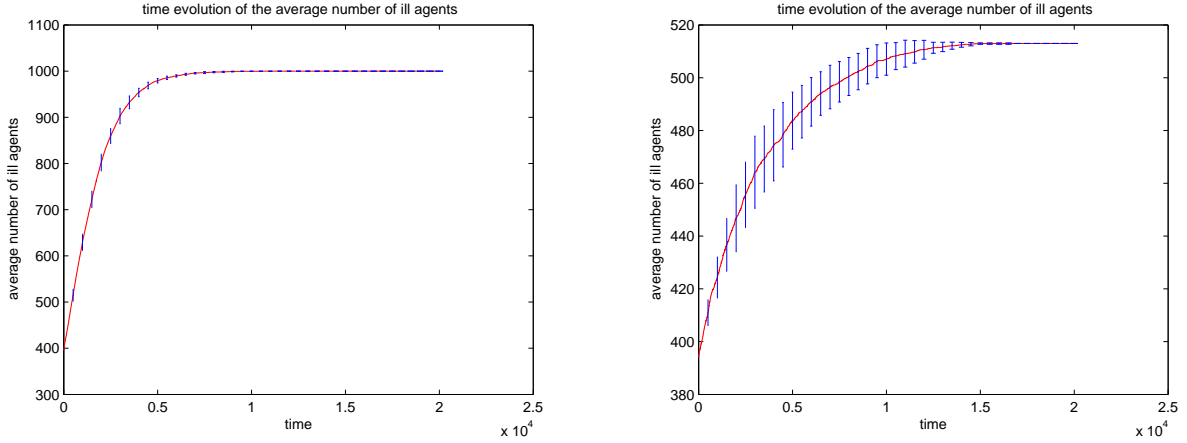


FIGURE 4.12 – Evolution du nombre moyen de malades pour le réseau *scale1* et les distributions *smalade4_1* et *smalade4_2*.

La figure 4.12 de gauche représente l'évolution du nombre moyen de malades lorsque la distribution *smalade4_1* est utilisée à chaque simulation. Nous constatons que cette distribution engendre bien une propagation totale de la maladie. De plus, une fois cette distribution fixée, les résultats ne se dispersent pratiquement plus autour de la moyenne. Les résultats des différentes simulations sont similaires.

La figure 4.12 de droite représente l'évolution du nombre moyen de malades lorsque la distribution *smalade4_2* est utilisée. Celle-ci engendre une propagation partielle de la maladie dans le réseau. De plus, nous constatons que les résultats des différentes simulations ne varient pratiquement plus, ils sont similaires.

Nous utiliserons ces deux distributions des malades lorsque nous travaillerons avec le réseau *scale1* et 40% de malades dans le réseau de départ.

Avec une proportion de malades de départ de 0.5

Considérons maintenant que la population soit composée de 50% de malades. Effectuons 15 simulations de propagation de la maladie dans le réseau avec, pour chaque simulation, une nouvelle distribution des malades. Nous obtenons

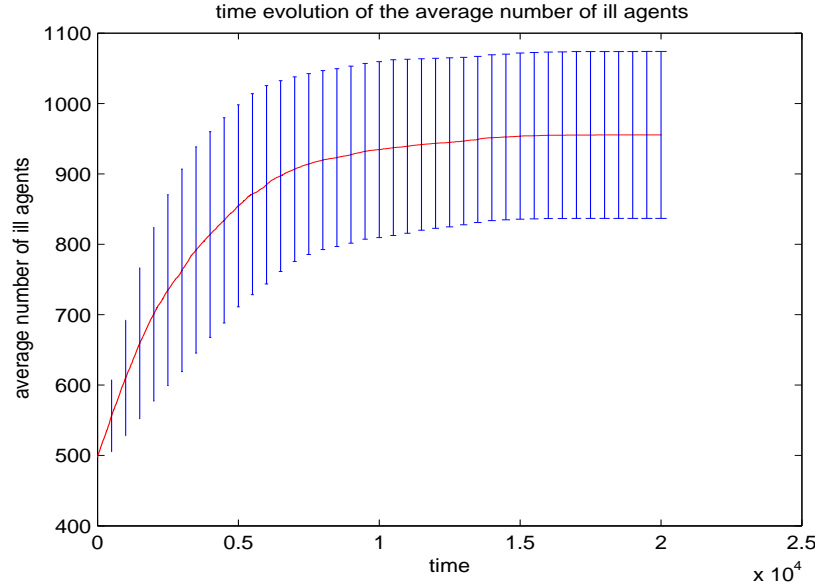


FIGURE 4.13 – Evolution du nombre moyen de malades pour le réseau *scale1* avec 50% de malades au départ.

Grâce à la figure 4.13, nous constatons que la distribution des malades a une influence sur la dynamique de la maladie étant donné la dispersion des résultats autour de la courbe moyenne. D'une simulation à l'autre, nous obtenons des résultats différents.

Pour ne plus avoir de si grandes différences, nous allons fixer des distributions. Lors de ces simulations, nous avons isolé deux distributions qui aboutissent à des situations différentes. La distribution *smalade5_1* engendre une propagation totale de la maladie dans le réseau tandis que *smalade5_2* engendre une propagation partielle.

Si nous effectuons 15 simulations de propagation de la maladie avec ces deux distributions, nous obtenons

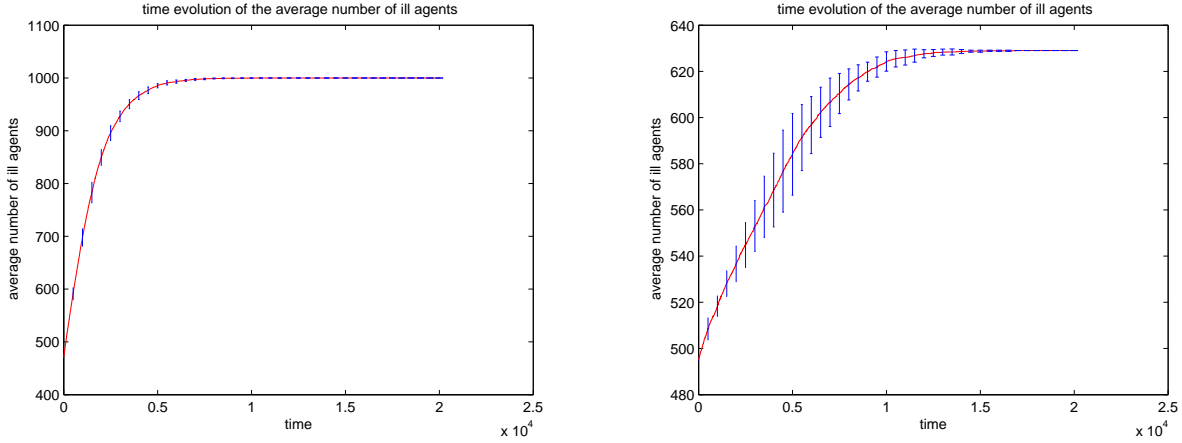


FIGURE 4.14 – Evolution du nombre moyen de malades pour le réseau *scale1* et les distributions *smalade5_1* et *smalade5_2*.

La figure 4.14 de gauche représente l'évolution du nombre moyen de malades dans le réseau lorsque la distribution *smalade5_1* est utilisée à chaque simulation. Nous constatons que cette distribution engendre une propagation totale de la maladie. De plus, une fois la distribution fixée, les résultats ne se dispersent pratiquement plus.

La figure 4.14 de droite, quant à elle, correspond à l'évolution du nombre moyen de malades dans le réseau lorsque la distribution *smalade5_2* est utilisée à chaque simulation.

Nous constatons que pour cette distribution, la propagation de la maladie est partielle. De plus, une fois la distribution fixée, les résultats ne se dispersent plus autant.

Nous utiliserons ces deux distributions des malades particulières lorsque nous travaillerons avec le réseau *scale1* et 50% de malades dans le réseau de départ.

4.2 Analyse du modèle de la propagation de l'information

Lorsque l'information se propage en plus dans le réseau, cela apporte une part de stochastique supplémentaire puisque la détermination des états d'information de chaque noeud se fait aléatoirement selon une proportion d'informés dans le réseau de départ. De plus, la propagation de l'information dans le réseau va s'effectuer selon une probabilité qu'une personne aura de devenir informée. Dans la suite de ce mémoire, nous avons choisi de fixer la détermination des états d'information pour l'ensemble des simulations que nous effectuerons.

Nous avons choisi une distribution d'informés telle que, aussi bien pour le réseau *random1*, que pour le réseau *scale1*, cette distribution engendre une propagation totale de l'information dans le réseau considéré. L'entièreté de la population sera informée après une certaine durée.

Nous avons choisi de fixer cette proportion d'informés puisque dans ce mémoire, nous nous intéressons à la propagation de la maladie lorsqu'il y a présence de l'information et non à la propagation de l'information dans un réseau. De ce fait, nous n'effectuerons pas d'analyse pour la propagation de l'information avec différentes proportions de personnes informées dans le réseau.

Parmi les distributions d'informés possibles, nous avons pu en isoler une, *info4*, dont 40% de la population sont informés.

Cette distribution *info4* engendre une propagation totale de l'information pour le cas de réseau *random1* et du réseau *Scale1*. En effet à partir de cette distribution d'informés et d'une probabilité³ de devenir informé $p_{info} = 0.6$, si nous effectuons 15 simulations avec le réseau *random1* et 15 pour le réseau *scale1*, nous obtenons

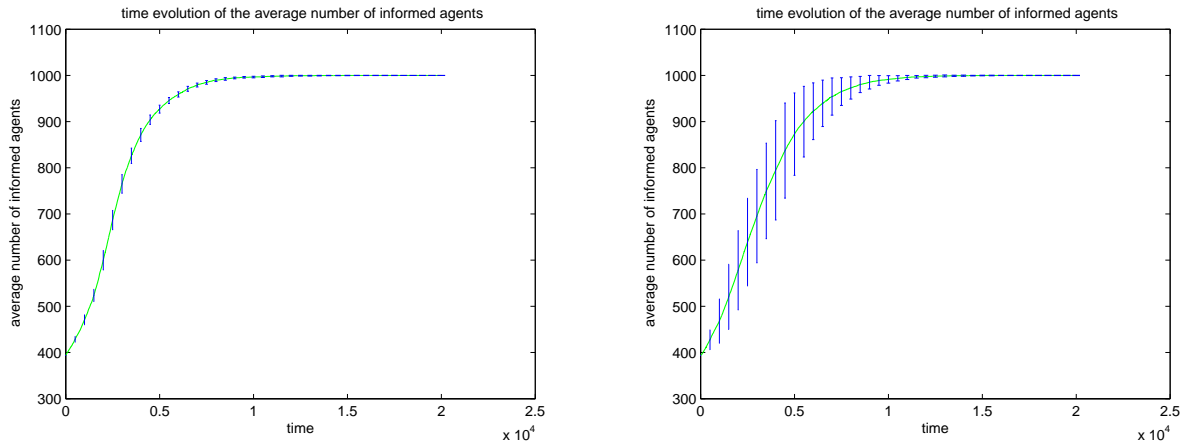


FIGURE 4.15 – Evolution du nombre moyen d'informés pour les réseaux *random1* et *scale1* et la distribution d'informés *info4*.

La figure 4.15 de gauche correspond à l'évolution de nombre moyen d'informés lorsque la distribution *info4* est utilisée à chaque simulation pour le réseau *random1*.

La figure 4.15 de droite correspond à l'évolution de nombre moyen d'informés lorsque la distribution *info4* est utilisée à chaque simulation pour le réseau *scale1*.

Nous constatons que la propagation de l'information est totale quelque soit le réseau *random1* ou *scale1* considéré.

Dans la suite, nous utiliserons cette distribution d'informés *info4* lorsque nous effectuerons les simulations avec l'information dans le réseau.

3. La valeur de la probabilité de devenir informé va influencer la vitesse de propagation de l'information mais n'influencera pas la dynamique de la propagation de celle-ci.

4.3 Evolution du nombre de malades en l'absence d'information dans le réseau

Dans cette section, nous nous intéressons à l'évolution du nombre de malades dans le réseau lorsqu'il n'y pas d'information. Nous étudierons l'évolution du nombre de malades pour les réseaux *random1* et *scale1* ainsi que l'évolution du nombre de malades ayant un certain degré entrant donné pour ces deux réseaux. Dans ce cas, nous ne considérerons que les noeuds qui ont ce degré entrant.

4.3.1 Random

Dans cette partie, nous considérons le réseau *random1* fixé. Nous allons effectuer l'étude de l'évolution du nombre de malades ainsi que celle du nombre de malades ayant un degré entrant k , pour plusieurs valeurs k différentes. Nous allons effectuer cette étude pour les différentes proportions de malades au départ dans le réseau. Pour cela, nous allons utiliser les distributions de malades que nous avons fixées dans la première section de ce chapitre. Pour la proportion de malades égale à 0.3, nous considérons les distributions *rmalade3_1* et *rmalade3_2*. Ces deux distributions aboutissent à des situations différentes. La première engendre une propagation totale de la maladie alors que la seconde engendre une propagation partielle.

Pour la proportion de malade égales à 0.4, une seule distribution *rmalade4* est utilisée, elle génère une situation où la propagation de la maladie est totale.

Il en est de même pour la distribution *rmalade5* lorsque la proportion de malades est égale à 0.5.

4.3.1.1 Evolution du nombre de malades

Nous allons nous intéresser à l'évolution du nombre de malades dans le réseau *random1* pour l'ensemble de ces distributions de malades de départ.

Nous utiliserons les valeurs $pmalade = 0.4$ et $pmalade = 0.8$ pour la probabilité de tomber malade.

Avec une proportion de malades au départ de 0.3

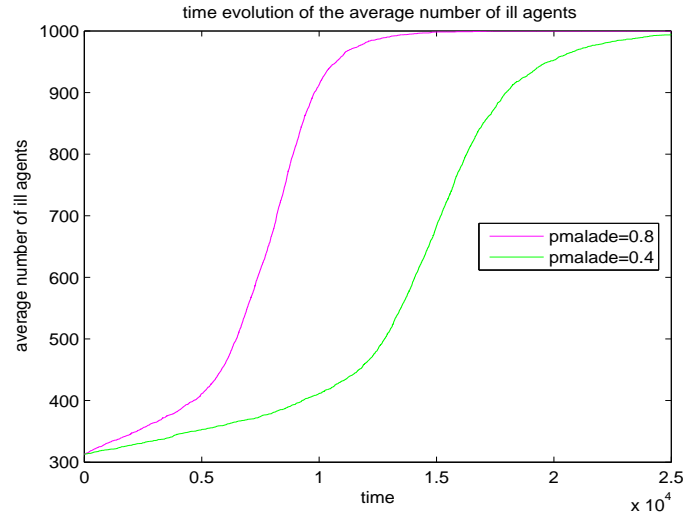


FIGURE 4.16 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade3_1*.

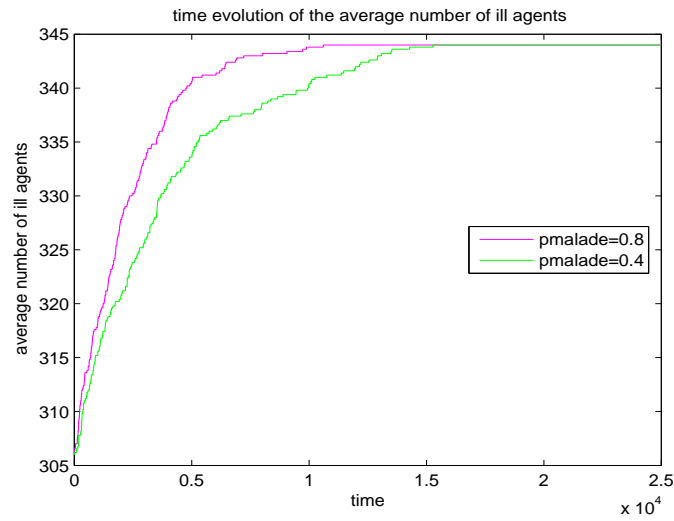


FIGURE 4.17 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade3_2*.

Avec une proportion de malades au départ de 0.4

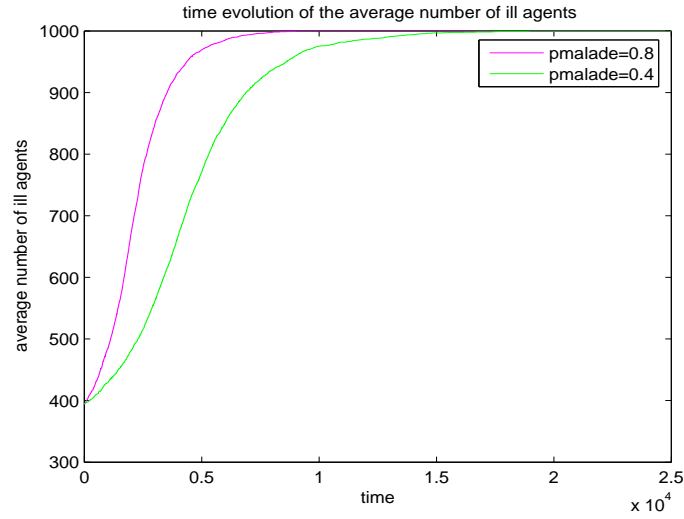


FIGURE 4.18 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade4*.

Avec une proportion de malades au départ de 0.5

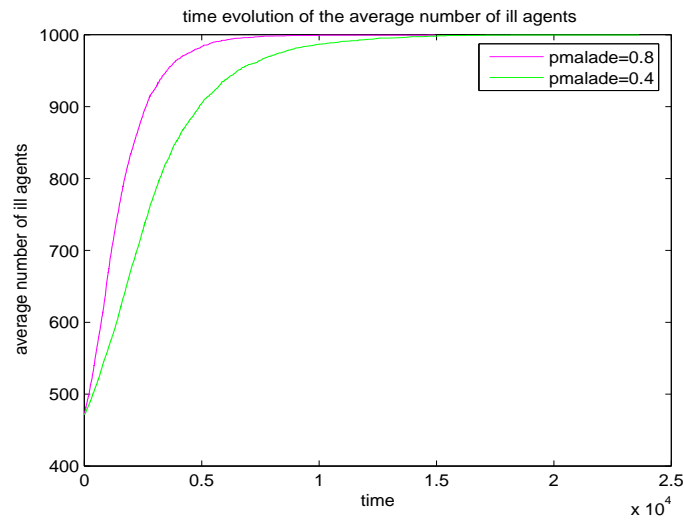


FIGURE 4.19 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade5*.

Résultats

Nous pouvons remarquer que la probabilité de tomber malade n'influence pas la valeur asymptotique du nombre de malades du réseau. En effet, nous constatons que pour chacune des distributions des malades considérées, le nombre asymptotique de malades dans le réseau sera le même quelque soit la probabilité de tomber malade considérée.

En fait, la probabilité de tomber malade va influencer la vitesse de propagation de la maladie. En effet, plus la probabilité de tomber malade va être importante, moins il faudra de temps pour que le nombre asymptotique de malades soit atteint.

Prenons par exemple la figure 4.17, nous constatons que pour $pmalade = 0.4$, il faut attendre environ 15000 étapes pour atteindre le nombre asymptotique de malades. Par contre, si nous considérons maintenant $pmalade = 0.8$, il ne faut plus que 10000 étapes pour que le même nombre de malades soit atteint.

Pour la figure 4.18, il faut attendre environ 19000 étapes pour que l'entièreté de la population soit infectée lorsque $pmalade = 0.4$ alors qu'avec $pmalade = 0.8$, il n'en faut plus qu'environ 9000.

Dans le réseau *random1* l'influence de la probabilité de tomber malade lors de la propagation de la maladie dans le réseau se porte sur la vitesse de propagation de celle-ci.

4.3.1.2 Evolution du nombre de malades ayant un degré entrant k

Nous allons maintenant étudier l'évolution du nombre de malades ayant un certain degré entrant k pour les différentes distributions de malades de départ. Nous travaillerons avec une probabilité de tomber malade $pmalade$ de 0.8. Nous venons de voir que cette probabilité n'influait pas le nombre asymptotique de malades du réseau *random1* mais la vitesse de propagation de la maladie. Le fait de choisir $pmalade = 0.8$ n'influera donc pas le nombre total de malades qu'il y aura dans le réseau après une longue période.

Nous allons effectuer cette étude pour plusieurs valeurs de degré entrant k différentes. Nous les avons choisies autour du degré moyen. Le degré moyen d'un réseau random se calcule par 2.4. Comme notre réseau *random1* est composé de 1000 noeuds et 10000 liens, nous avons

$$\langle k \rangle = \frac{2m}{n} = \frac{20000}{1000} = 20$$

Nous allons prendre comme valeur pour k , les valeurs 15, 20 et 25, ainsi nous aurons l'évolution pour le degré moyen et deux autres valeurs autour de ce degré moyen.

Nous allons utiliser les différentes distributions des malades fixées dans la première section de ce chapitre.

Avec une proportion de malades au départ de 0.3

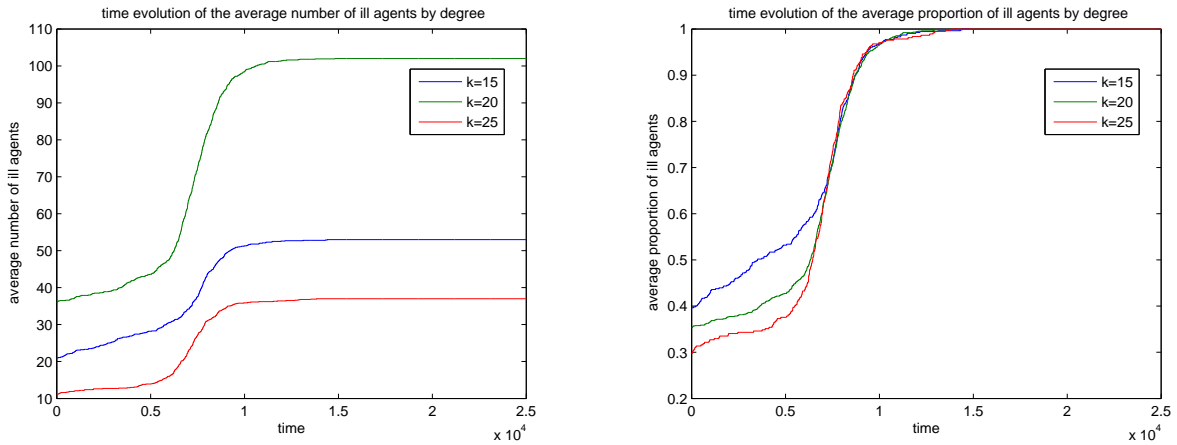


FIGURE 4.20 – Evolution du nombre moyen et de la proportion moyenne de malades ayant un degré entrant k pour le réseau *random1* et la distribution des malades *rmalade3_1*.

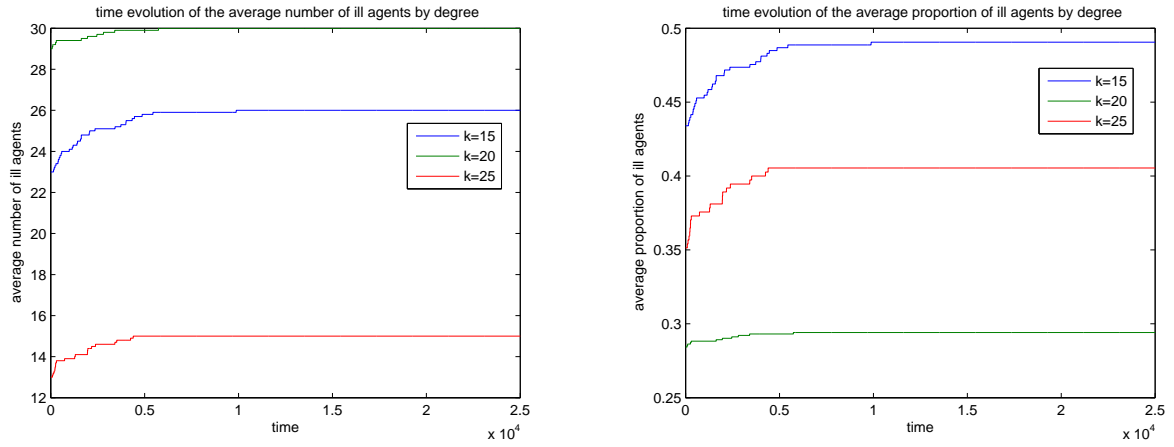


FIGURE 4.21 – Evolution du nombre moyen et de la proportion moyenne de malades ayant un degré entrant k pour le réseau *random1* et avec la distribution des malades *rmalade3_2*.

Avec une proportion de malades au départ de 0.4

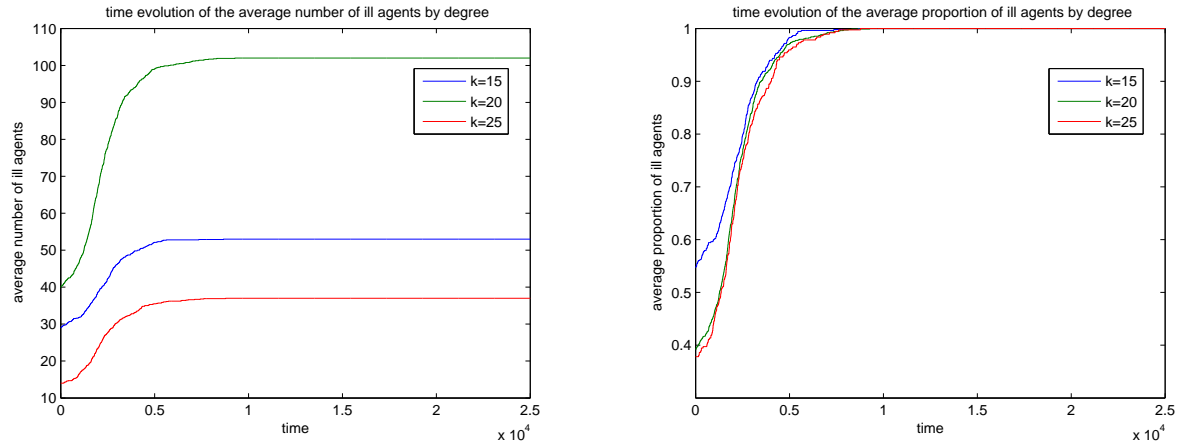


FIGURE 4.22 – Evolution du nombre moyen et de la proportion moyenne de malades ayant un degré entrant k pour le réseau *random1* et avec la distribution des malades *rmalade4*.

Avec une proportion de malades au départ de 0.5

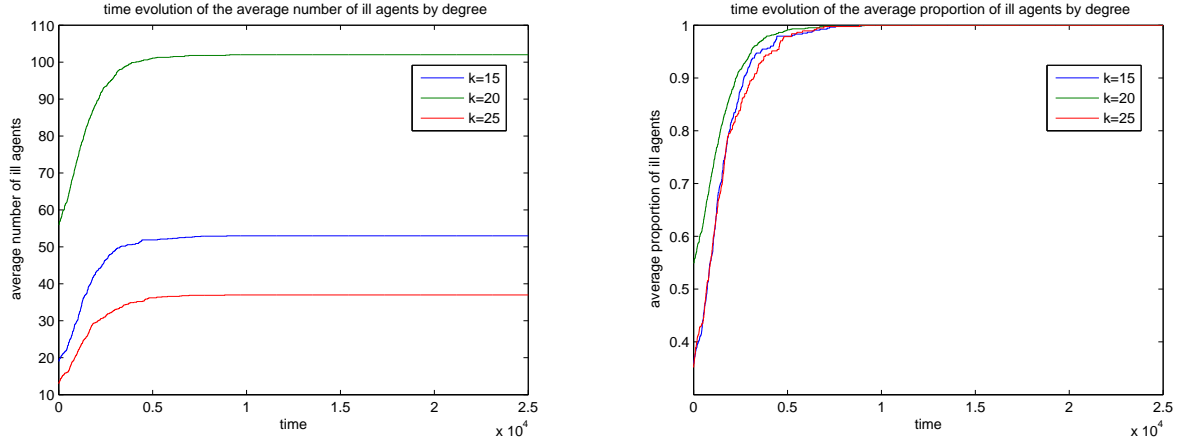


FIGURE 4.23 – Evolution du nombre moyen et de la proportion moyenne de malades ayant un degré entrant k pour le réseau *random1* et avec la distribution des malades *rmalade5*.

Résultats

Dans le cas du réseau *random1*, pour chaque distribution considérée, nous constatons qu'il y a plus de noeuds malades de degré entrant $k = 20$ que les autres degrés considérés. Comme $k = 20$ correspond au degré moyen du réseau, il est évident qu'il y aura plus de noeuds de ce degré dans le réseau. S'il y a plus de noeuds de degré $k = 20$, il y aura donc plus de noeuds de ce degré infectés que des noeuds infectés des autres degrés considérés.

Considérons les figures 4.20, 4.22 et 4.23. Dans ces cas, la propagation de la maladie est totale. Si nous considérons l'évolution de la proportion moyenne de malades ayant un certain degré, avec les valeurs $k = 15$, $k = 20$ et $k = 25$, nous constatons que les évolutions des proportions de malades ayant un certain degré semblent être identiques pour les trois degrés considérés. Nous en déduisons, que la propagation de la maladie est donc uniforme. Elle ne touche pas plus rapidement un degré particulier qu'un autre.

Lorsque la propagation de la maladie n'est pas totale, seulement une partie des noeuds de chaque degré sera touchée. Ceci est évident, puisque quand la propagation est partielle, seule une part de la population est touchée.

Nous constatons grâce à la figure 4.21, que la proportion moyenne asymptotique de malades de chaque degré n'atteint même pas 0.5.

4.3.2 Scale-Free

Nous considérons maintenant le réseau *scale1*. Nous allons effectuer l'étude de l'évolution du nombre de malades ainsi que celle du nombre de malades ayant un degré entrant k , pour plusieurs valeurs k différentes.

Nous allons utiliser les distributions des malades que nous avons dans la première section de ce chapitre. Pour la proportion de malades égale à 0.3, nous considérons les distributions *smalade3_1* et *smalade3_2*. Ces deux distributions aboutissent à des situations différentes. La première engendre une propagation totale de la maladie alors que la seconde engendre une propagation partielle.

Pour la proportion de malades égale à 0.4, nous considérons les distributions *smalade4_1* et *smalade4_2* dont la première amène à une propagation totale de la maladie et la deuxième, une propagation partielle.

Enfin nous considérons *rmalade5_1* et *rmalade5_2* lorsque la proportion de malades est égale à 0.5. La première distribution aboutit à une propagation totale de la maladie alors que la deuxième aboutit à une propagation partielle.

4.3.2.1 Evolution du nombre de malades

Nous allons étudier l'évolution du nombre de malades dans le réseau *scale1* pour l'ensemble de ces distributions de malades de départ. Nous utiliserons les valeurs $pmalade = 0.4$ et $pmalade = 0.8$.

Avec une proportion de malades au départ de 0.3

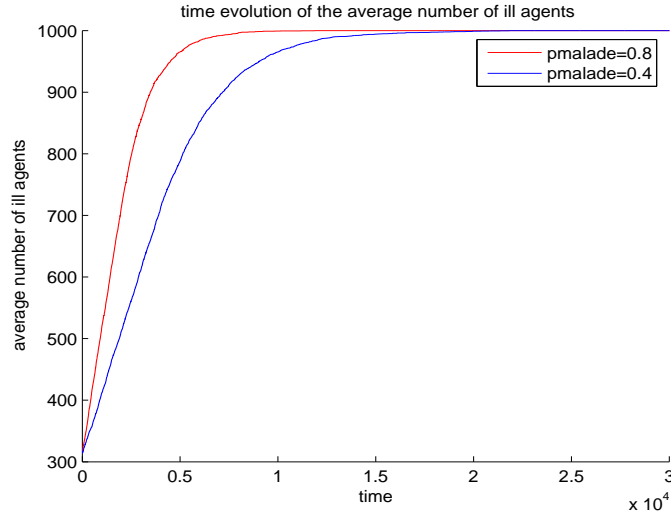


FIGURE 4.24 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade3_1*.

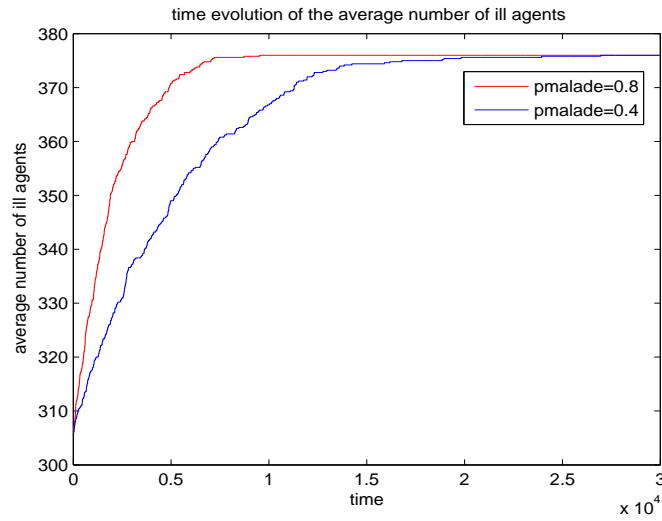


FIGURE 4.25 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade3_2*.

Avec une proportion de malades au départ de 0.4

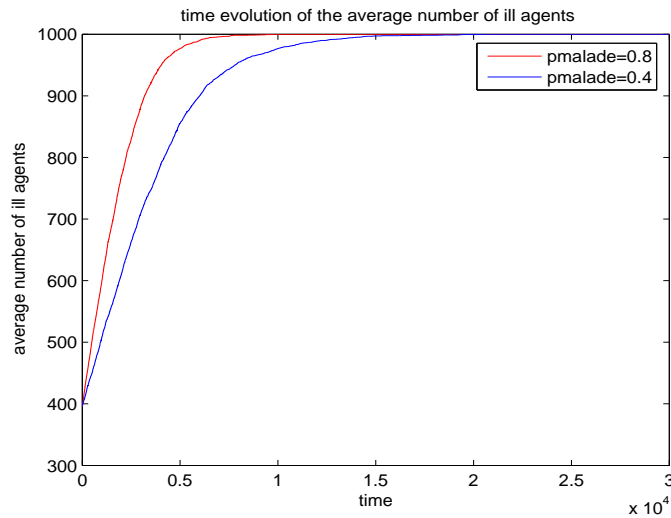


FIGURE 4.26 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade4_1*.

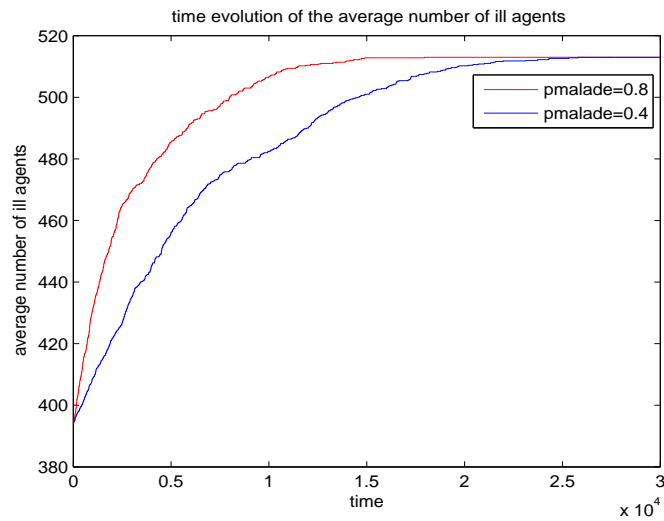


FIGURE 4.27 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade4_2*.

Avec une proportion de malades au départ de 0.5

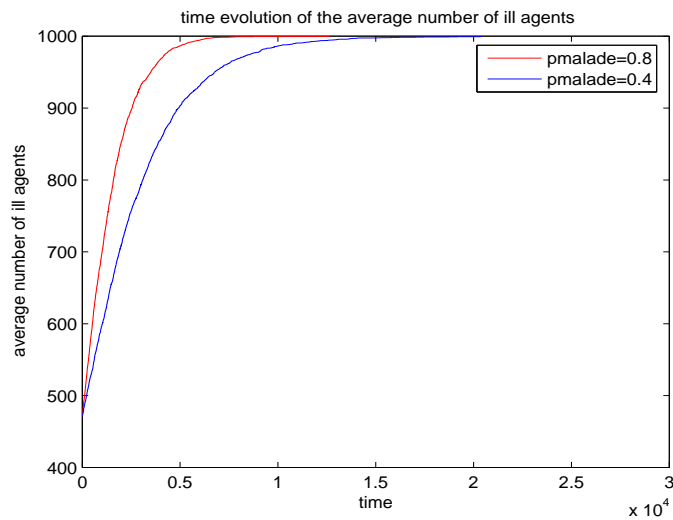


FIGURE 4.28 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade5_1*.

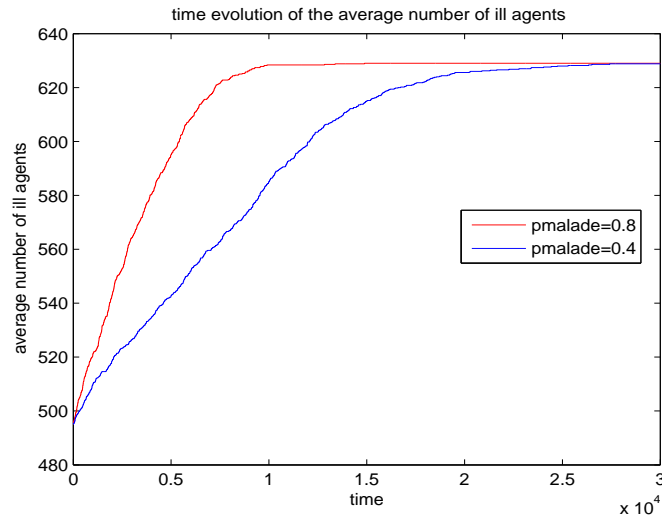


FIGURE 4.29 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade5_2*.

Résultats

Comme avec le réseau *random1*, nous constatons que pour le réseau *scale1* la valeur de la probabilité de tomber malade n'influence pas le nombre asymptotique de malades et ce quelque soit la distribution des malades considérée.

Par contre, la probabilité de tomber malade influence la vitesse de propagation de la maladie. Nous pouvons constater, grâce aux figures ci-dessus, qu'avec une probabilité plus élevée, et ce quelque soit la proportion de malades au départ dans le réseau, le temps nécessaire pour que le nombre asymptotique de malades soit atteint sera plus court qu'avec une probabilité plus faible.

4.3.2.2 Evolution du nombre de malades ayant un degré entrant k

Nous allons maintenant étudier l'évolution du nombre de malades ayant un certain degré entrant k pour différentes distributions des malades de départ. Nous travaillerons avec une probabilité de tomber malade $pmalade$ de 0.8. Nous avons vu que cette probabilité de tomber malade n'influencait pas le nombre asymptotique de malades du réseau mais la vitesse de propagation de la maladie. Le fait de choisir une probabilité égale à 0.8 n'influence donc pas la dynamique de la maladie.

Nous allons effectuer cette étude pour plusieurs valeurs k différentes. Elles ont été choisies en fonction de la distribution des degrés entrant parmi les noeuds du réseau *scale1*.

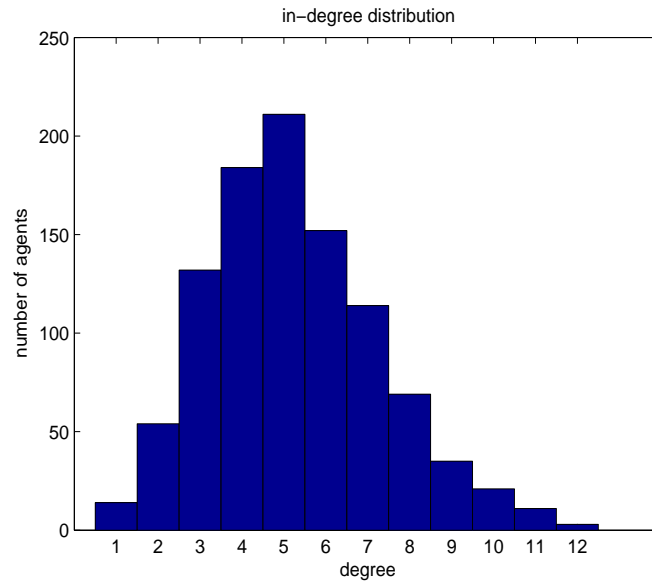


FIGURE 4.30 – Histogramme du nombre de noeuds par degré entrant pour le réseau *scale1*.

La figure 4.30 représente la distribution des degrés entrant pour le réseau *scale1*. Au vu de cette distribution, nous choisissons de travailler avec les valeurs $k = 4$, $k = 8$ et $k = 12$.

Nous allons utiliser les différentes distributions des malades concernant le réseau *scale1* que nous avons fixées.

Avec une proportion de malades au départ de 0.3

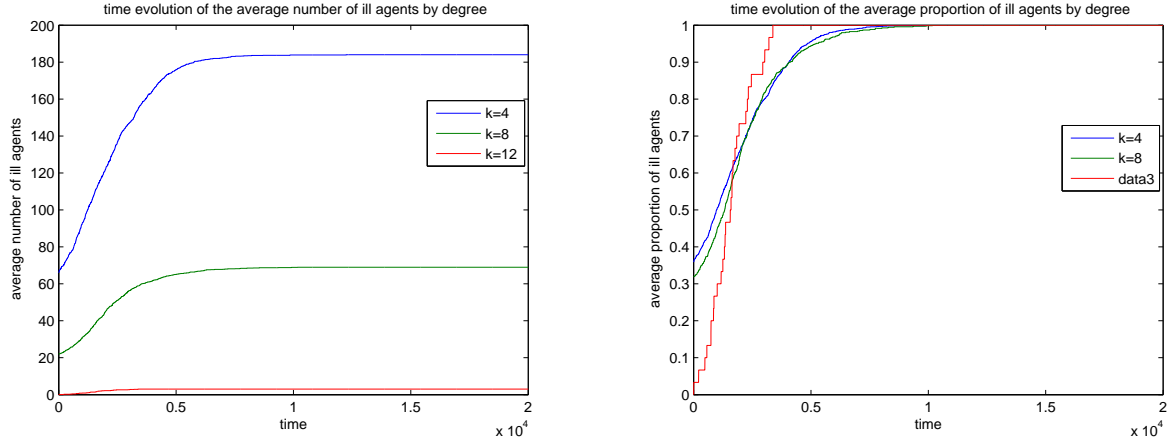


FIGURE 4.31 – Evolution du nombre moyen et de la proportion moyenne du nombre de malades ayant un degré entrant k pour le réseau *scale1* et la distribution des malades *smalade3_1*.

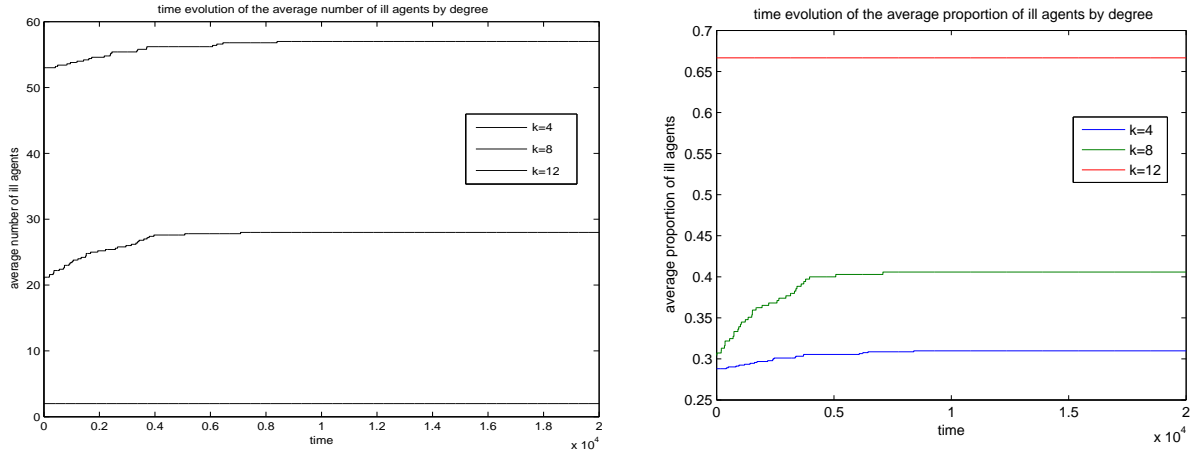


FIGURE 4.32 – Evolution du nombre moyen et de la proportion moyenne du nombre de malades ayant un degré entrant k pour le réseau *scale1* la distribution des malades *smalade3_2*.

Avec une proportion de malades au départ de 0.4

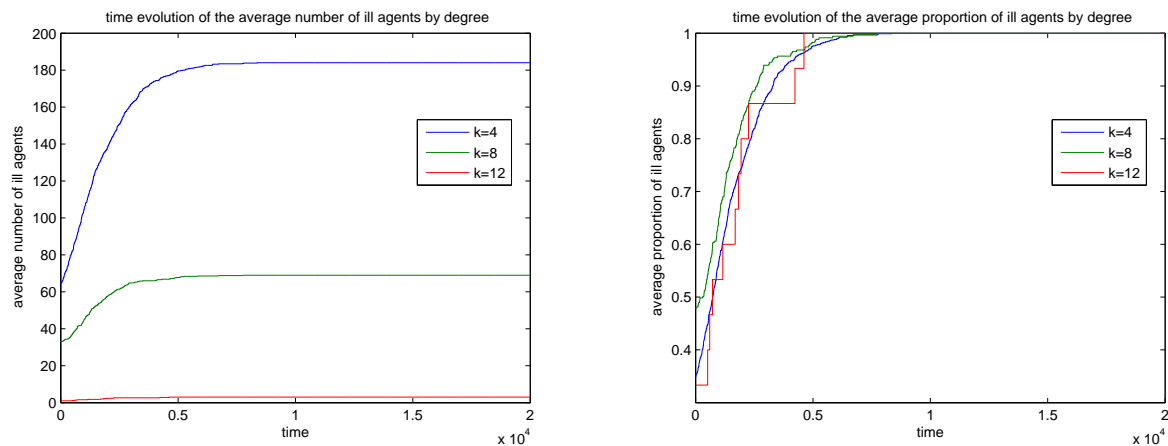


FIGURE 4.33 – Evolution du nombre moyen et de la proportion moyenne du nombre de malades ayant un degré entrant k pour le réseau *scale1* la distribution des malades *smalade4_1*.

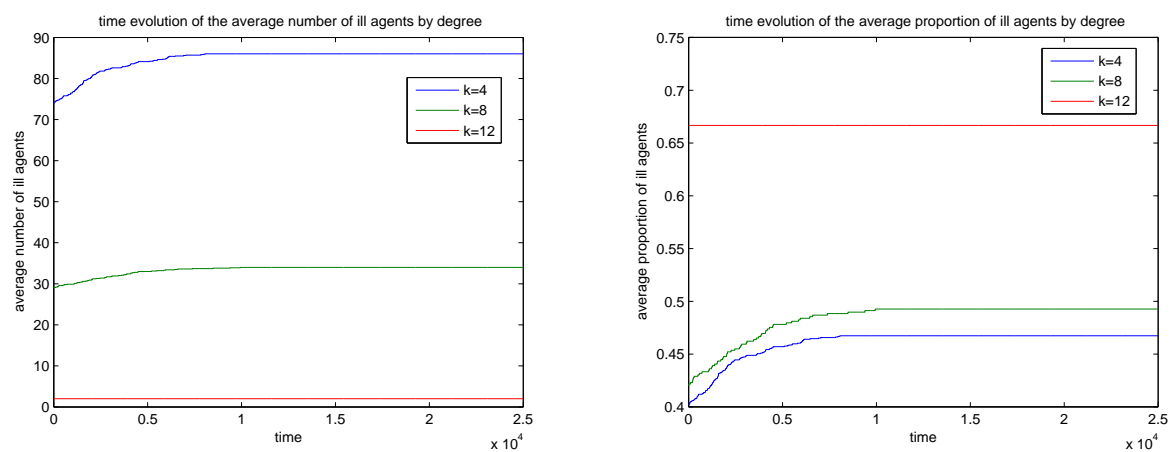


FIGURE 4.34 – Evolution du nombre moyen et de la proportion moyenne du nombre de malades ayant un degré entrant k pour le réseau *scale1* la distribution des malades *smalade4_2*.

Avec une proportion de malades au départ de 0.5

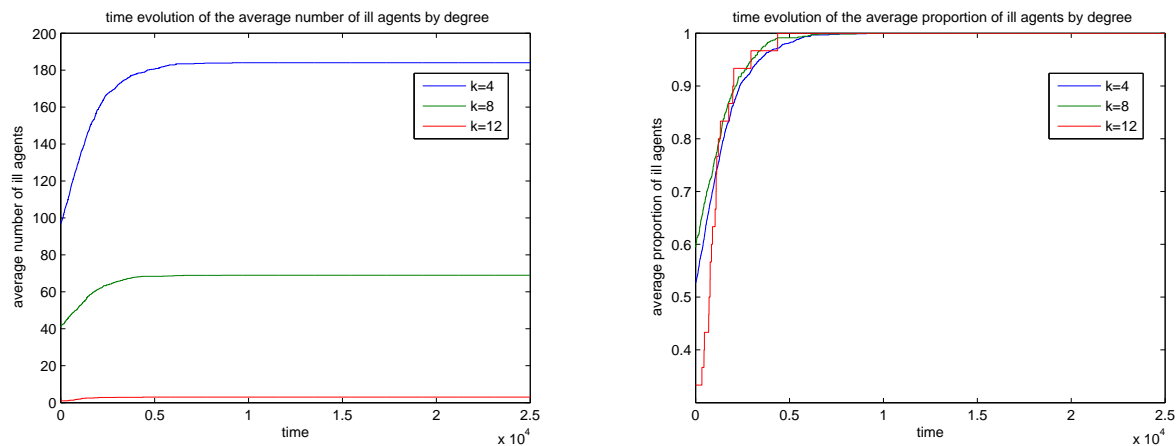


FIGURE 4.35 – Evolution du nombre moyen et de la proportion moyenne du nombre de malades ayant un degré entrant k pour le réseau *scale1* la distribution des malades *smalade5_1*.

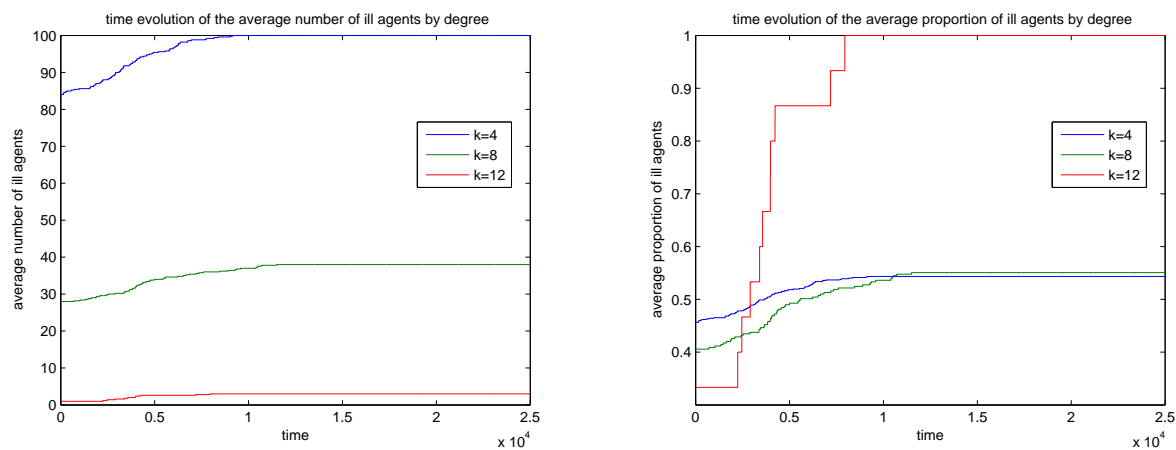


FIGURE 4.36 – Evolution du nombre moyen et de la proportion moyenne du nombre de malades ayant un degré entrant k pour le réseau *scale1* la distribution des malades *smalade5_2*.

Résultats

Nous pouvons constater, grâce aux figures ci-dessus, qu'il y a plus de noeuds infectés de degré entrant $k = 4$ au départ. Si on considère la distribution des degrés entrant (cfr figure 4.30), la part de noeuds de degré entrant $k = 4$ est déjà plus importante. Il est alors évident que lors de la détermination des états de santé, qu'une part plus importante de noeud de degré entrant $k = 4$ sera infectée par rapport aux autres valeurs de degré considérées.

Nous constatons aussi que, lorsque la propagation est totale (cfr figures 4.31, 4.33 et 4.35), les évolutions de la proportion moyenne de malades pour chaque degré sont similaires pour les trois valeurs de k considérées. La propagation de la maladie semble donc être uniforme, lorsque celle-ci est totale. La part du nombre de malades de chaque degré entrant évoluent de la même façon avec le temps.

Lorsque la propagation de la maladie n'est pas totale, le nombre de malades de chaque degré évolue peu. Il est possible néanmoins, comme sur la figure 5.36, que l'ensemble des noeuds d'un certain degré soit malade. Ceci s'explique par le fait que la population compte peu de noeuds de ce degré et que justement ce peu de noeuds est touché par la maladie au bout d'un certain temps.

4.4 Evolution du nombre de malades en présence d'information dans le réseau

Dans cette section, nous nous intéressons à l'évolution du nombre de malades dans le réseau lorsque l'information y est présente.

Nous étudierons l'évolution du nombre de malades avec la distribution d'informés *info4* déterminée début du chapitre. Elle correspond à une distribution dont 40% de la population est informée.

Nous allons étudier l'évolution du nombre de malades pour différentes valeurs de la probabilité de devenir informé. Ces valeurs sont $pinfo=0$, $pinfo=0.4$, $pinfo=0.6$ et $pinfo=0.8$.

L'étude de l'évolution du nombre de malades dans le réseau avec $pinfo=0$ correspond à l'étude de l'évolution du nombre de malades lorsque l'information ne se propage pas dans le réseau, mais est tout de même présente. En effet, nous avons supposé que 40% de la population sont informés. Une fois que cette probabilité est supérieure à 0, l'information se propage dans le réseau. Au fur et à mesure du temps, il y aura plus de personnes informées dans le réseau.

Nous avons dû fixer les probabilités qu'aura un noeud de tomber malade : si ce noeud est non-informé, sa probabilité de tomber malade sera de $pmalade = 0.8$, alors que s'il est informé, elle sera $pmalade = 0.4$.

4.4.1 Random

Dans cette partie, nous considérons le réseau *random1*. Nous allons effectuer l'étude de l'évolution des malades pour les différentes valeurs de *pinfo* et pour différentes proportions de malades au départ dans le réseau. Pour cela, nous allons utiliser les distributions de malades que nous avons fixées. Ces distributions sont *rmalade3_1*, *rmalade3_2*, *rmalade4* et *rmalade5*.

Avec une proportion de malades au départ de 0.3

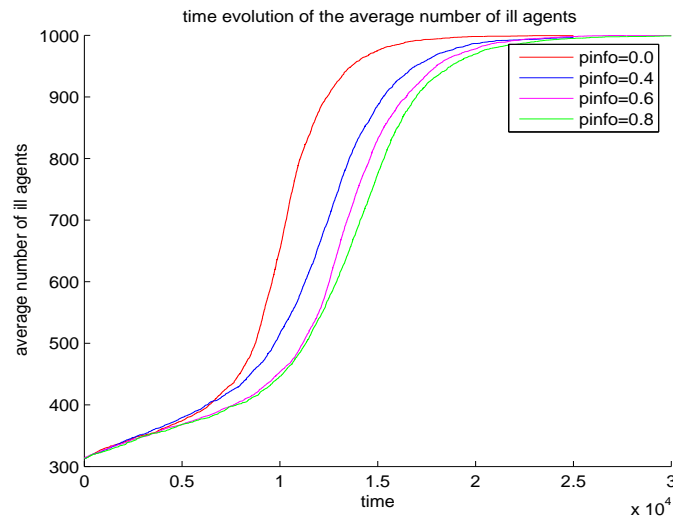


FIGURE 4.37 – Evolution du nombre moyen de malades lorsqu'il y a présence d'information pour le réseau *random1* et la distribution *rmalade3_1*.

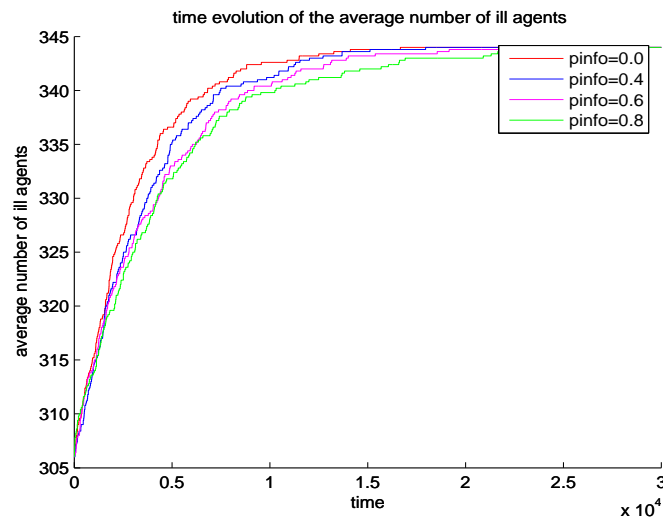


FIGURE 4.38 – Evolution du nombre moyen de malades lorsqu'il y a présence d'information pour le réseau *random1* et la distribution *rmalade3_1*.

Avec une proportion de malades au départ de 0.4

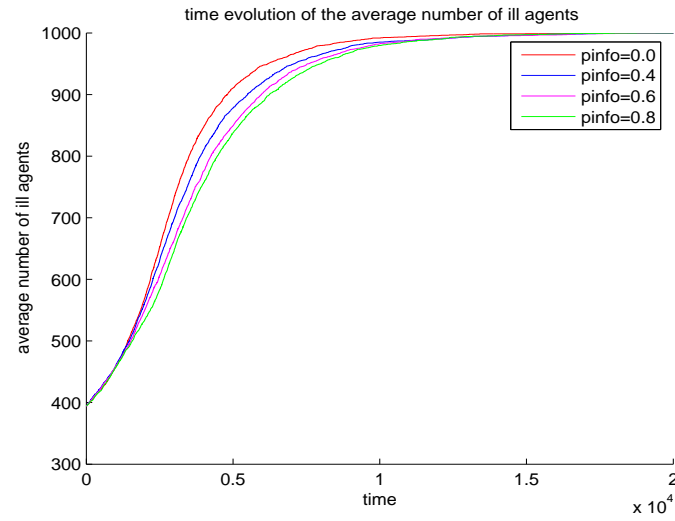


FIGURE 4.39 – Evolution du nombre moyen de malades lorsqu'il y a présence d'information pour le réseau *random1* et la distribution *rmalade4*.

Avec une proportion de malades au départ de 0.5

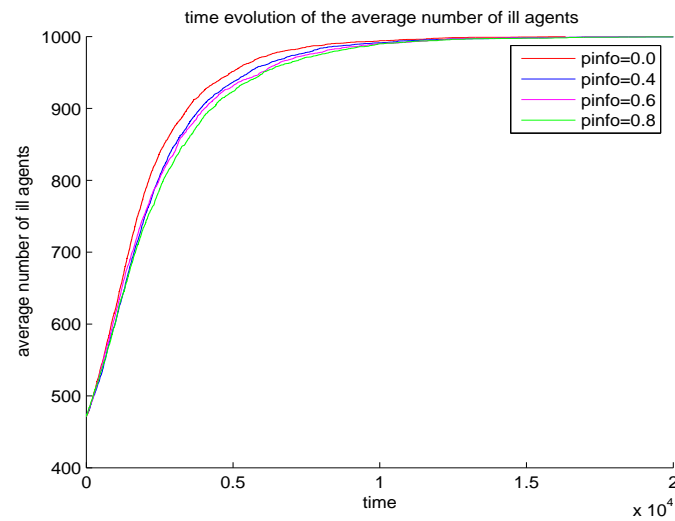


FIGURE 4.40 – Evolution du nombre moyen de malades lorsqu'il y a présence d'information pour le réseau *random1* et la distribution *rmalade5*.

Résultats

Pour le réseau *random1*, la présence de l'information ne modifie pas la dynamique de la maladie. En effet, les distributions *rmalade3_1*, *rmalade4* et *rmalade5* engendrent toujours la propagation totale de la maladie alors que la distribution *rmalade5* engendre la propagation partielle de la maladie dans le réseau.

Nous remarquons aussi, quelque soit la distribution considérée, que plus la probabilité de devenir informé est grande, plus la vitesse de propagation de la maladie est lente. Plus l'information sera transmissible, plus il faudra du temps pour que la maladie se propage.

4.4.2 Scale-Free

Nous considérons maintenant le réseau *scale1*. Nous allons effectuer l'étude de l'évolution des malades pour les différentes valeurs de *pinfo* et pour différentes proportions de malades au départ dans le réseau. Nous utilisons les différentes distributions de malades *smalade3_1*, *smalade3_2*, *smalade4_1*, *smalade4_2*, *smalade5_1* et *rmalade5_2*.

Avec une proportion de malades au départ de 0.3

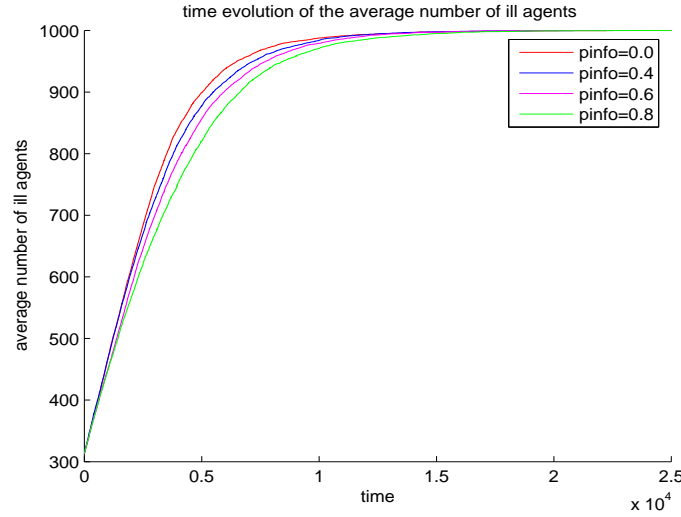


FIGURE 4.41 – Evolution du nombre moyen de malades lorsqu'il y a présence d'information pour le réseau *scale1* et la distribution *smalade3_1*.

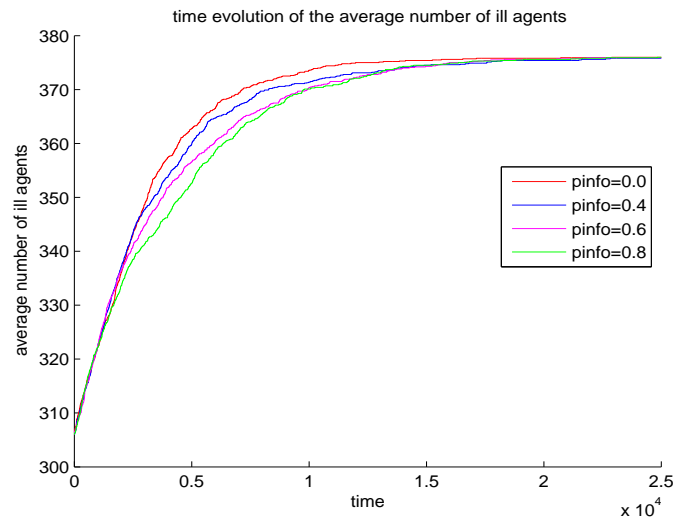


FIGURE 4.42 – Evolution du nombre moyen de malades lorsqu’il y a présence d’information pour le réseau *scale1* et la distribution *smalade3_2*.

Avec une proportion de malades au départ de 0.4

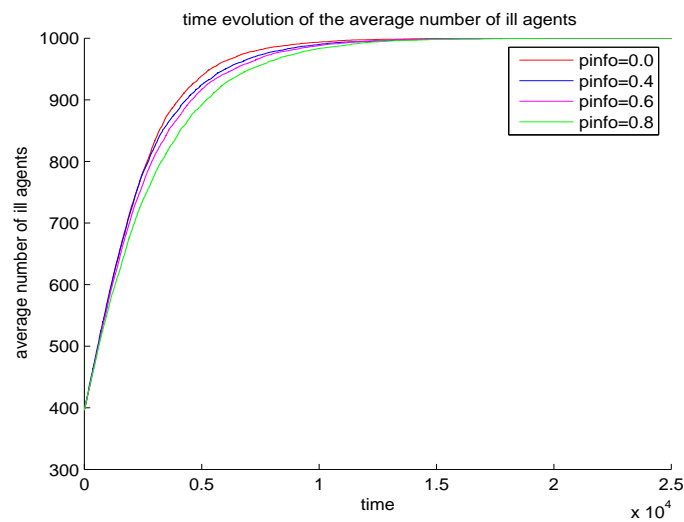


FIGURE 4.43 – Evolution du nombre moyen de malades lorsqu’il y a présence d’information pour le réseau *scale1* et la distribution *smalade4_1*.

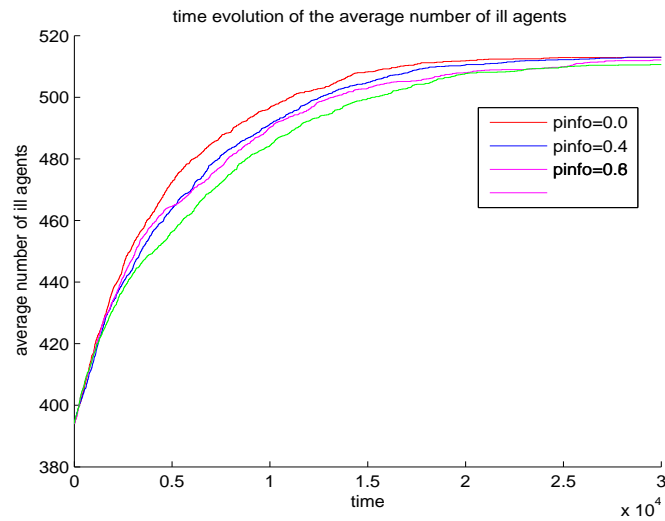


FIGURE 4.44 – Evolution du nombre moyen de malades lorsqu’il y a présence d’information pour le réseau *scale1* et la distribution *smalade4_2*.

Avec une proportion de malades au départ de 0.5

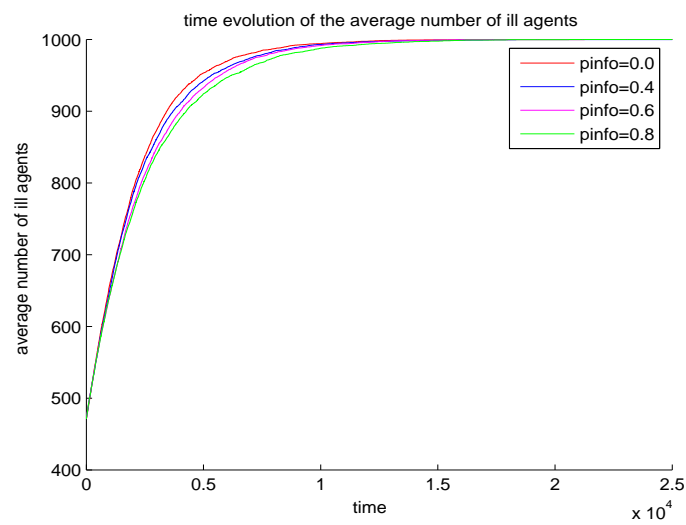


FIGURE 4.45 – Evolution du nombre moyen de malades lorsqu’il y a présence d’information pour le réseau *scale1* et la distribution *smalade5_1*.

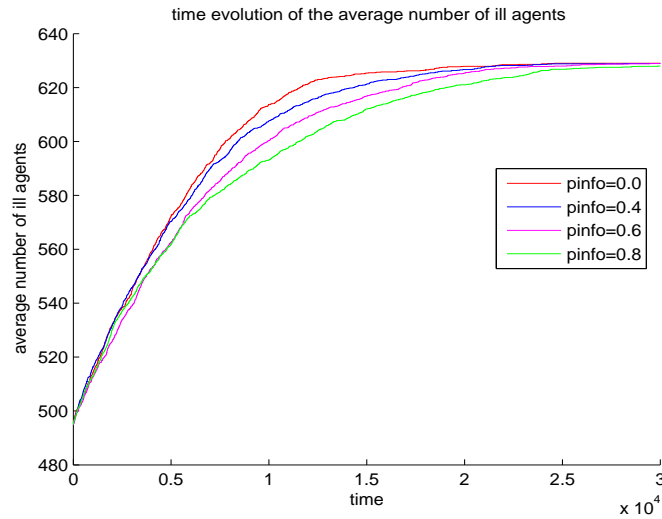


FIGURE 4.46 – Evolution du nombre moyen de malades lorsqu’il y a présence d’information pour le réseau *scale1* et la distribution *smalade5_2*.

Résultats

Pour le réseau *scale1*, la présence d’information dans le réseau ne modifie pas le nombre asymptotique des malades pour les distributions considérées. En effet, les distributions qui engendraient la propagation totale de la maladie lorsqu’il n’y avait pas d’information, engendrent toujours la propagation totale de la maladie même si l’information est présente dans le réseau. Nous remarquons aussi, grace aux figures ci-dessus, que quelque soit la distribution de malades considérée, plus la probabilité de devenir informé est élevée, plus la vitesse de propagation de la maladie est lente.

Plus l’information a une probabilité de se transmettre élevée, plus il faudra du temps pour que la maladie se propage.

Chapitre 5

Comparaison de la propagation de la maladie avec et sans information

5.1 Introduction

Dans ce chapitre, nous allons comparer l'évolution du nombre de malades lorsque l'information circule, ou non, dans le réseau. Nous allons effectuer ces comparaisons pour les deux réseaux considérés, *random1* qui est un réseau de type Random et *scale1* qui est un réseau de type Scale-Free.

Nous considérerons chaque distribution de malades que nous avons fixée au chapitre 4 pour chacun des deux réseaux et nous propagerons la maladie dans le réseau, tantôt sans information et tantôt avec.

Pour ces simulations, nous considérons que la probabilité qu'a un noeud de tomber malade s'il n'est pas informé est $pmalade = 0.8$. Par contre, si celui-ci est informé, sa probabilité de tomber malade ne sera plus que de $pmalade = 0.4$. Nous avons montré au chapitre précédent que la probabilité de devenir malade influençait la vitesse de propagation de la maladie. En choisissant ces probabilités plutôt que d'autres, nous n'influencerons pas le nombre total de personnes malades dans le réseau.

Nous allons effectuer l'étude de l'évolution du nombre de malades pour plusieurs situations. La première correspond à la situation où aucune information n'est présente dans le réseau.

Ensuite nous introduirons l'information. Pour ce faire, nous considérerons la distribution des personnes informées fixée au chapitre 4, *info4*. Cette distribution implique que 40% de la population soient informés.

Nous procéderons tout d'abord à la propagation de la maladie dans ce réseau avec l'information qui ne se propage pas ($pinfo=0$). La proportion d'informés au départ est de 0.4 et le reste jusqu'à la fin de la propagation.

Ensuite, nous augmenterons la probabilité de devenir informé. Nous utiliserons les valeurs $info=0.4$ et $pinfo=0.8$.

5.2 Comparaison pour le réseau de type Random

Nous allons considérer dans cette partie le réseau *random1*. Nous allons effectuer l'ensemble des simulations de propagation de la maladie avec chaque distribution de malades que nous avons fixée pour le réseau. Ces distributions sont *rmalade3_1*, *rmalade3_2*, *rmalade4* et *rmalade5*.

Avec une proportion de malades au départ de 0.3

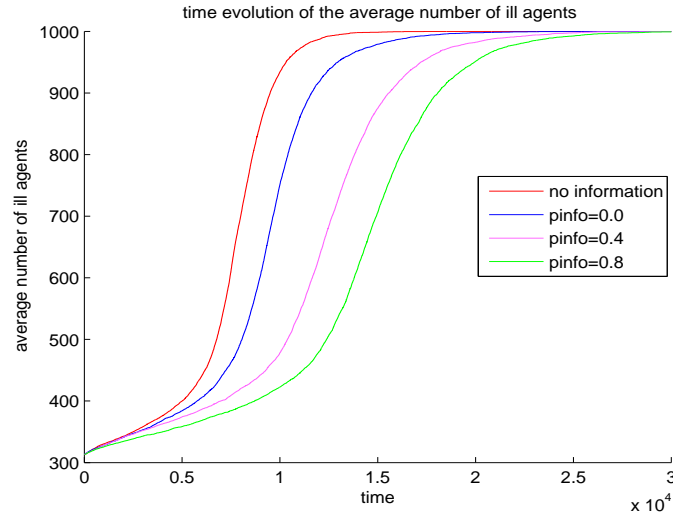


FIGURE 5.1 – Evolution du nombre moyen de malades pour la distribution *rmalade3_1*.

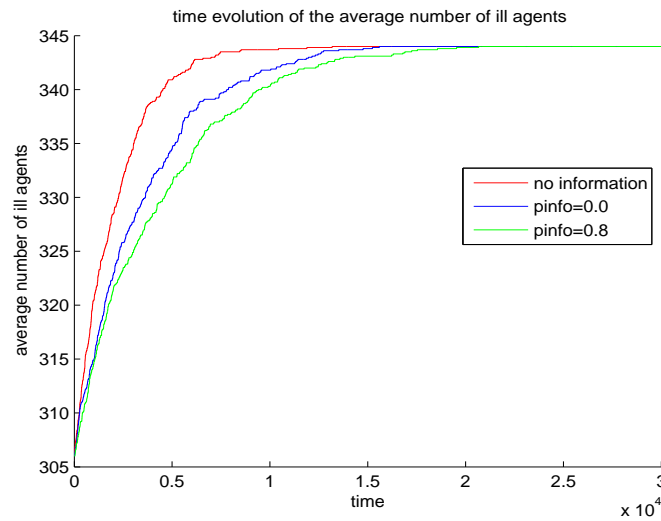


FIGURE 5.2 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade3_2*.

Avec une proportion de malades au départ de 0.4

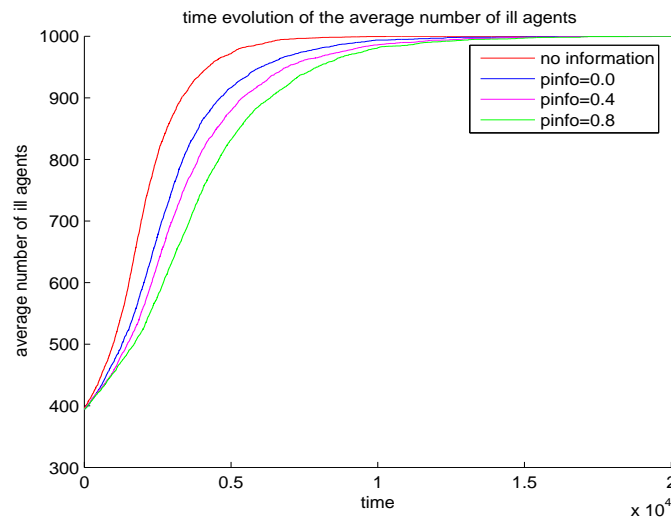


FIGURE 5.3 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade4*.

Avec une proportion de malades au départ de 0.5

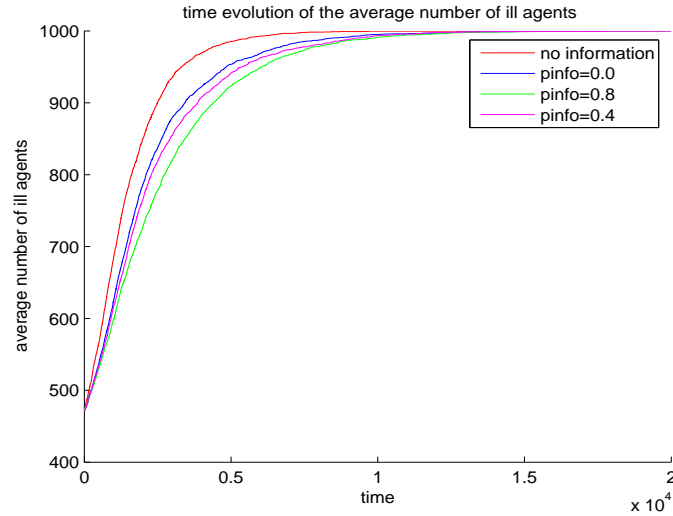


FIGURE 5.4 – Evolution du nombre moyen de malades pour le réseau *random1* et la distribution *rmalade5*.

Résultats

Nous pouvons constater que, pour le réseau *random1* de type Random, la présence de l'information dans le réseau, même sans propagation de celle-ci, réduit la vitesse de la propagation de la maladie.

En effet, quelque soit la distribution de malades considérée, nous constatons déjà une différence considérable entre la courbe représentant l'évolution du nombre de malades lorsqu'il n'y a aucune information (courbe rouge) et la courbe représentant l'évolution du nombre de malades lorsque certaines personnes¹ sont informées et que la probabilité de devenir informé est 0 (courbe bleue).

Nous constatons aussi que plus la probabilité de devenir informé est grande, plus la différence entre les courbes d'évolution du nombre de malades est marquée. En effet, plus l'information aura une grande capacité de se diffuser, plus la vitesse de propagation de la maladie est réduite. La présence d'information dans le réseau *random1* a donc un impact positif sur la propagation de la maladie.

1. Nous avons considéré que lorsqu'il y a information, 40% de la population est informé au départ.

5.3 Comparaison pour le réseau de type Scale-Free

Nous allons maintenant considérer dans cette partie le réseau *scale1*. Nous allons effectuer l'ensemble des simulations de propagation de la maladie avec chaque distribution de malades que nous avons fixées pour ce réseau. Ces distributions sont *smalade3_1*, *smalade3_2*, *smalade4_1*, *smalade4_2*, *smalade5_1* et *smalade5_2*.

Avec une proportion de malades au départ de 0.3

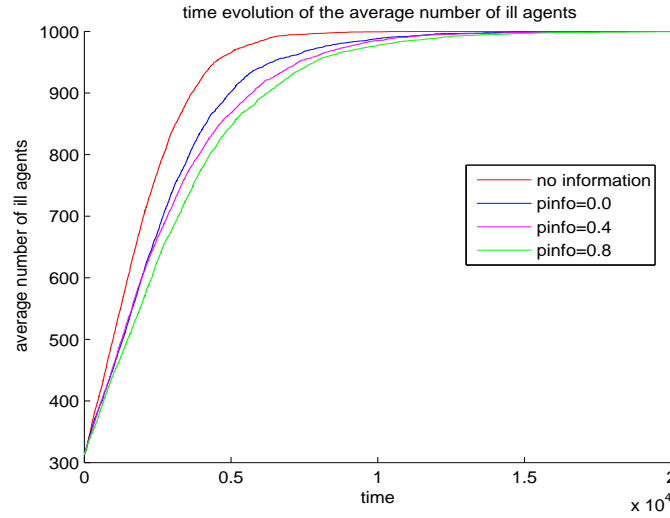


FIGURE 5.5 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade3_1*.

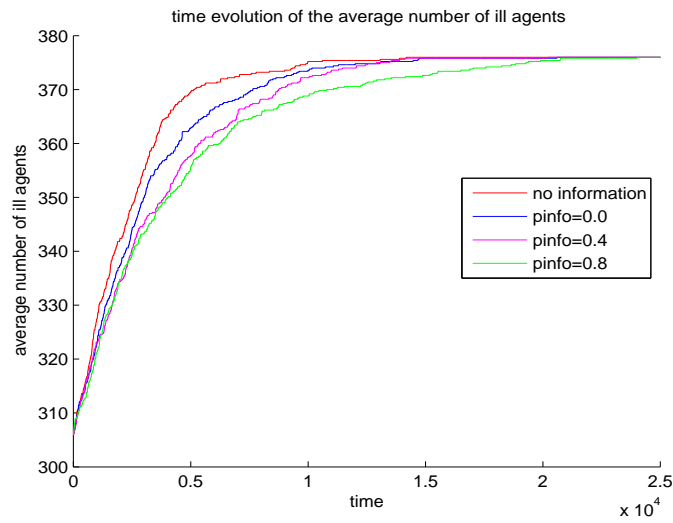


FIGURE 5.6 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade3_2*.

Avec une proportion de malades au départ de 0.4

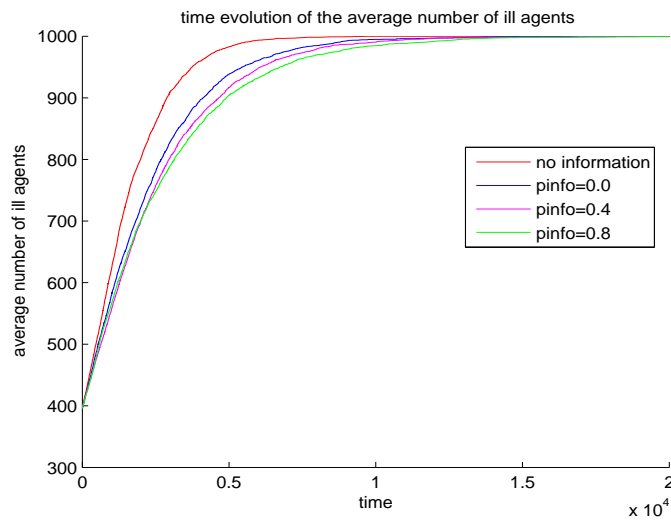


FIGURE 5.7 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade4_1*.

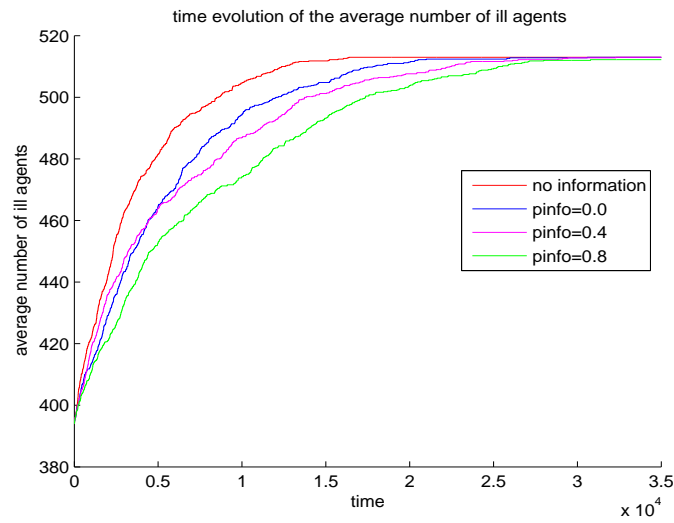


FIGURE 5.8 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade4_2*.

Avec une proportion de malades au départ de 0.5

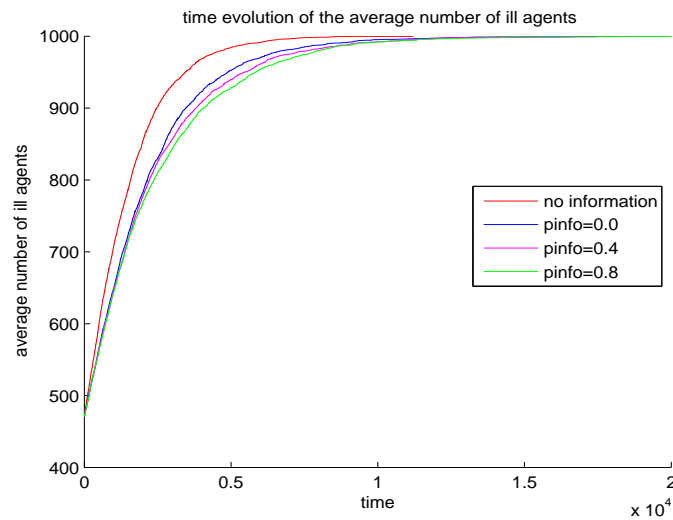


FIGURE 5.9 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade5_1*.

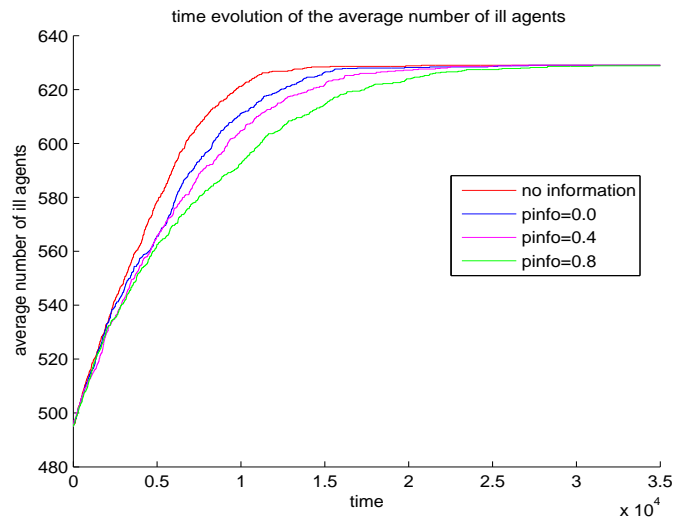


FIGURE 5.10 – Evolution du nombre moyen de malades pour le réseau *scale1* et la distribution *smalade5_2*.

Résultats

Pour le cas du réseau *scale1*, nous obtenons des résultats similaires à ceux obtenus pour le cas du réseau *random1*.

Dès qu'il y a présence d'information dans le réseau, même si cette information ne se diffuse pas, nous pouvons constater une différence considérable. La maladie propage plus rapidement lorsqu'aucune information n'est présente dans le réseau.

Nous constatons aussi que plus la probabilité de devenir informés est grande.

Dans le réseau *scale1* la présence de l'information joue un rôle bénéfique lors de la propagation de la maladie.

Conclusion

Nous pouvons affirmer que la présence de l'information joue un rôle bénéfique lors de la propagation de la maladie quelque soit le réseau considéré, *random1* et *scale1*. En effet, nous avons choisi de fixer les réseaux *random1* et *scale1* afin de travailler avec le même réseau pour l'ensemble des simulations effectuées mais si nous avions choisi un autre réseau de type Random ou de type Scale-Free, nous serions arrivés aux mêmes résultats.

Il ne faut donc pas négliger les campagnes d'information lors d'épidémies, car celles-ci ont un rôle important à jouer. Plus la population sera avertie des précautions à prendre, plus la maladie se transmettra lentement.

Dans ce mémoire, nous avons dû faire certaines hypothèses lors de la modélisation des réseaux.

Par exemple, nous avons fait l'hypothèse de travailler avec des réseaux statiques. Les connexions que nous avons choisies pour chaque noeud, sont restées identiques tout au long de la propagation de la maladie. Nous pourrions, par exemple imaginer, que le réseau puisse évoluer à chaque étape de temps. Un noeud apprenant qu'un de ses amis est malade, va couper cette connection à risque et en créera peut-être une autre avec un nouvel ami, qui lui est sain.

Nous avons aussi effectué des hypothèses sur la propagation de la maladie et de l'information. Nous avons choisi le schéma de propagation suivant : pour qu'un noeud devienne malade ,ou informé², il faut que plus de la moitié de ses noeuds parents soient eux-mêmes malades, ou informés. A chaque instant, nous choisissons une personne et regardions si plus de la moitié de ses parents était malade, ou informée.

Ce schéma n'est évidemment pas unique. Nous aurions pu imaginer que la maladie se transmettait lors de relations sexuelles. Dans ce cas, pour une personne en particulier, nous n'aurions plus dû vérifier si la moitié des connections étaient infectées, parce qu'à priori, nous n'avons pas de relations sexuelles avec tous nos amis. Nous aurions plutôt choisi une connection particulière, avec qui la personne serait supposée avoir des relations sexuelles. Cette personne pourra donc être infectée si la personne avec qui elle a des rapports l'est.

2. Nous avons utilisé le même schéma de propagation pour la maladie et l'information.

Nous pourrions encore imaginer d'autres pistes. Au lieu de travailler avec le modèle SI comme nous l'avons fait dans ce mémoire, nous aurions pu considérer :

- soit le modèle SIS où une fois guéri de l'infection, un noeud serait de nouveau susceptible de devenir malade.
- soit le modèle SIR où l'état *recovery* signifierait la mort d'un noeud suite à l'infection. Dans ce cas, le réseau n'est plus statique puisqu'à tout moment des noeuds peuvent mourir, dans ce cas les connections de ces noeuds sont supprimées.

Ces pistes ne sont évidemment pas les seules, nous pourrions encore en trouver d'autres. Elles pourraient même faire l'objet d'un mémoire futur ou même d'une thèse.

Bibliographie

- [1] Guido Galdarelli, *Scale-Free Networks : Complex webs in nature and technology*, Editions OXFORD University Press, 2007
- [2] Fernando Vega-Redondo, *Complex Social Networks*, Editions CAMBRIDGE University Press, 2007
- [3] A. Barrat, M. Barthélemy, A. Vespignami, *Dynamical Processes on complex Networks*, Editions CAMBRIDGE University Press, 2008
- [4] Duncan J. Watts, *Six Degrees, The science of a connected age*, Editions NORTON, 2004
- [5] Duncan J. Watts, *Small Worlds, The dynamics of Networks between Order and Randomness*, Editions PRINCETON University Press, 2004
- [6] Linked, The New-Sciences of Networks, *Alber-Laslo Barabasi*, Editions PERSEUS, Cambrigde, Massachusetts, 2002
- [7] Frank Callier, *Théorie des graphes*, FUNDP, 2006-2007
- [8] Jean-Charles Delvenne, *Théorie des graphes*, FUNDP, 2009-2010,

Annexe

Cette annexe reprend les fichiers MATLAB qui ont servis tout au long de ce mémoire. Ils correspondent à

- deux fonctions utilisées pour créer des réseaux de type Random et de type Scale-Free
- un fichier utilisé lors de la propagation de la maladie
- deux fonctions écrites par Timoteo Carletti pour évaluer le diamètre d'un réseau.

Voici le code correspondant à la fonction qui crée un réseau de type Random.

```
%Fonction qui genere un reseau de type random de parametre n et m
%Elle renvoie le reseau noeud cree et le vecteur de degres sortants

function [noeud,l] = random(n,m)

noeud=cell(n,1); %Creation d'un tableau de cellules de dimension nx1

p=2*m/(n*(n-1)); %Determination de la probabilite
                %qu'a un noeud d'etre connecte avec un autre

l=binornd(n-1,p,[1 n]); %Determiner du degre sortant de chaque noeud
                        %en fonction de la probabilite 'p'

for i=1:n %Pour l'ensemble des noeud du reseau
    if l(i)>0 %si le noeud i a au moins une connection
        while length(noeud{i})<l(i) %et tant que nous n'avons pas atteint
                                    %le nombre de connections
                                    %de ce noeud i
            h=floor(rand()*n+1); %nous selectionnons un noeud du reseau
            if h~=i %si ce noeud est different du noeud i
                noeud{i}=union(noeud{i},h); %nous l'ajoutons à la liste
                                           %de connections de noeud i
            end
        end
    end
end
end
```

Le code correspondant à la fonction qui crée des réseaux de type Scl-Free est :

```
%Fonction qui genere un reseau de type scale free de parametre n et gamma
%Elle renvoie le reseau noeud cree et le vecteur de degres sortants

function [noeud,l] = scale(n,gamma)

noeud=cell(n,1); %Creation d'un tableau de cellules de dimension nx1

%Determination du coefficient de normalisation  $C^{\gamma}$ 
C=0;
for i=1:n-1
    C=C+(1/(igamma));
end
C=1/C; %C correspond au coefficient de normalisation

%Determination du vecteur des probabilités cumulées
somme=C;
F=somme;
for i=2:n-1
    somme=somme+(C/igamma);
    F=[F somme]; %F est le vecteur des probabilites cumulees
end

r=rand(n,1); %selection d'un nombre aléatoire entre 0 et 1
           %pour chaque noeud du réseau

for i=1:n
    v=find(F<r(i)); %v contiendra l'emplacement des éléments de F
                  %qui verifiront la cond <r(i).
    l(i)=length(v)+1; %le degre sortant d'un noeud correspond au nombre
                    %d'elements de F qui verifie la condition <r(i)
                    %incrementé de 1
end

for i=1:n %Pour chaque noeud du réseu
    if l(i)>0 %si le noeud i a au moins une connection
        while length(noeud{i})<l(i) %et tant que nous avons pas atteint
            %le nombre de connections de ce noeud i
```

```

        h=floor(rand()*n+1); %nous sélectionnons un noeud du reseau
        if h~=i %si ce noeud est different du noeud i
            noeud{i}=union(noeud{i},h); %nous l'ajoutons à la liste
                                   %de connections de noeud i
        end
    end
end
end
end

```

Voici le code correspondant au fichier qui est utilisé lors de l'étude de la propagation de la maladie en présence, ou non, de l'information dans le réseau.

```

%Ce fichier est utilise lors de l'étude de la propagation de la maladie dans
%le réseau.
%Le programme effectue 'niter' simulations et enregistre l'evolution du nombre
%des malades et du nombre d'informés.
%Ensuite, il effectue une moyenne des différentes simulations.

```

```

niter=15; %nombre de simulations
tmax=35000; %nombre d'etape de temps

```

```

noeud=random1;
%noeud=scale1;
n=length(noeud);

```

```

parent=cell(n,1); %Determination du liste parentale du reseau
for i=1:n          %Pour chaque noeud i du reseau
    for j=1:n
        if (length(find(noeud{j}==i))>0) %si le noeud i
                                           %fait partie de la liste des
                                           %connections du noeud j
            parent{i}=union(parent{i},j); %alors le noeud j est un noeud
                                           %parent du noeud i
        end
    end
    l2(i)=length(parent{i}); %Determination du degre entrant de
                             %chaque noeud
end

```

```

ks=[15,20,25]
%ks=[4,8,12]; % the program will compute the number of ill agents whose degrees are in ks

```

```

pmalade=0.8; %probabilite de tomber malade si non-informé

```

```

pinfo=0.4; %probabilité de devenir informé
pmaladeinfo=0.4; %probabilité de tomber malade si informé

numbmal=zeros(niter,tmax);
    %tableau des nombres de malades à chaque étape de temps pour l'ensemble
    %des simulations
avnumbmal=zeros(1,tmax);
    %tableau des nombres moyens de malades à chaque étape de temps
stdnumbmal=zeros(1,tmax);
    %tableau des déviations standards du nombre moyen de malades à chaque
    %étape de temps
numbinfo=zeros(niter,tmax);
    %tableau des nombres d'informés à chaque étape de temps pour l'ensemble
    %des simulations
avnumbinfo=zeros(1,tmax);
    %tableau des nombres moyens d'informés à chaque étape de temps
stdnumbinfo=zeros(1,tmax);
    %tableau des déviations standards du nombre moyen d'informés

numbmalk=zeros(niter,tmax,length(ks));
    %tableau des nombres de malades de degré k à chaque étape de temps et
    %pour l'ensemble
    %des simulations
propbmalk=zeros(niter,tmax,length(ks));
    %tableau des proportions de malades de degré k à chaque étape de temps et
    %pour l'ensemble
    %des simulations
avnumbmalk=zeros(length(ks),tmax);
    %tableau des nombres moyen de malades de degré k à chaque étape de temps
avpropbmalk=zeros(length(ks),tmax);
    %tableau des proportions moyennes de malades de degré k à chaque étape de temps
stdnumbmalk=zeros(length(ks),tmax);
    %tableau des déviation standant du nombre moyen de malades de degré k

tic
for s=1:niter, %Pour chaque simulation,
    s

    malade=malade5_6; %La distribution malade5_6 est considérée
                    %Pour changer de distribution, inscrire le nom de la
                    %distribution à la plade de 'malade5_6'

    info=info4_1; %L'information est présente dans le réseau
    %info=binornd(1,0.0,[1,n]); %L'information n'est pas présente dans le

```

```

                                %réseau. (PS: ne pas oublier de mettre
                                %'pinfo' à 0 si nous ne travaillons pas
                                %avec l'information

numbmals(s,1)=sum(malade); %Comptage du nombre de malades au temps 1 et
                                %lors de la simulation s
numbinfo(s,1)=sum(info); %Comptage du nombre d'informés au temps 1 et
                                %lors de la simulation s

for j=1:length(ks), %Comptage du nombre de malades de degré k
                                %pour chaque degré considéré
    kk=ks(j); %Pour chaque degré kk considéré
    aa=find(l2==kk); %Nous considérons la liste des degrés entrants
                                %égal au degré kk
    numbmalk(s,1,j)=sum(malade(aa)); %Nous comptabilisons le nombre de
                                %malades ayant ce degré kk
    propbmalk(s,1,j)=sum(malade(aa))/length(aa);
                                %Nous calculons la proportion de malades de degré au
                                %temps 1 et lors de la simulation s
end

for t=2:tmax %Pour chaque étape de temps

    i=floor(rand*n + 1); %Nous sélectionnons un noeud

    %Propagation de l'information:
    if (info(i)==0) %Si cet agent est non-informé,
        %possibilité que son état d'information évolue
        parentinfo=sum(info(parent{i})); %Comptage du nombre de noeuds
                                %parents informés
        if parentinfo>l2(i)/2 %Si il y a plus d'informés
            %que de non-informés
            info(i)=binornd(1,pinfo); %Le noeud peut devenir informé
                                %avec une probabilité 'pinfo'
        end
    end

    %Propagation de la maladie
    if (malade(i)==0) %Si ce noeud est sain
        %possibilité que son état de santé évolue
        parentmalade=sum(malade(parent{i})); %Comptage du nombre de
                                %noeuds parents informés
        if parentmalade>l2(i)/2 %si il y a plus de malades que de sains
            if info(i)==0 %et que le noeud i est non-info

```

```

        malade(i)=binornd(1,pmalade); %Le noeud peut devenir
                                   %malade avec une
                                   %probabilité 'pmalade'
    else                            %Par contre si le noeud i est info
        malade(i)=binornd(1,pmaladeinfo);
                                   %Le noeud peut devenir malade avec une
                                   %probabilité 'pmaladeinfo'
    end
end
end

numbmals(s,t)=sum(malade); %Comptage du nombre de malades
                        %au temps t et lors de la simulation s
numbinfo(s,t)=sum(info); %Comptage du nombre d'informés
                        %au temps t et lors de la simulation s

for jj=1:length(ks) %Comptage du nombre de malades de degré k
                    %pour chaque degré considéré
    kk=ks(jj);
    aa=find(l2==kk);
    numbmalk(s,t,jj)=sum(malade(aa));
    propbmalk(s,t,jj)=sum(malade(aa))/length(aa);
end

end %[boucle de temps]

end %[boucle des simulations]
toc

avnumbmals=mean(numbmals); %Evaluation du nombre moyen de malades à chaque
                        %étape de temps
stdnumbmals=std(numbmals); %Evaluation des déviations standards pour les
                        %nombres de malades à chaque étape de temps

avnumbinfo=mean(numbinfo); %Evaluation du nombre moyen d'informés à chaque
                        %étape de temps

stdnumbinfo=std(numbinfo); %Evaluation des déviations standards pour les
                        %nombres d'informés à chaque étape de temps

for jj=1:length(ks),
    avnumbmalk(jj,:)=mean(numbmalk(:,:,jj));
    %Evaluation du nombre moyen de malades de degré k à chaque étape
    %de temps et pour tous les degrés considérés
    avpropbmalk(jj,:)=mean(propbmalk(:,:,jj));
    %Evaluation de la proportion moyenne de malades de degré k à chaque

```



```

        %etape de temps et pour tous les degrés considérés
        stdnumbmalk(jj,:)=std(numbmalk(:, :,jj));
        %Evaluation des déviations standards pour les nombres d'informés
        %à chaque étape de temps
    end

%Représentation de l'évolution du nombre moyen de malades avec barres
%d'erreur
figure
plot(1:tmax,avnumbmalk,'r')
hold on
errorbar(1,avnumbmalk(1),stdnumbmalk(1))
for i=1:tmax
    if mod(i,500)==0
        hold on
        errorbar(i,avnumbmalk(i),stdnumbmalk(i))
    end
end
title('time evolution of the average number of ill agents')
xlabel('time');
ylabel('average number of ill agents');
legend('pmalade=0.8')
%legend('no information')

%Représentation de l'évolution du nombre moyen d'informés avec barres
%d'erreur
figure
plot(1:tmax,avnumbinfo,'g')
hold on
errorbar(1,avnumbinfo(1),stdnumbinfo(1))
for i=1:tmax
    if mod(i,500)==0
        hold on
        errorbar(i,avnumbinfo(i),stdnumbinfo(i))
    end
end
title('time evolution of the average number of informed agents')
xlabel('time');
ylabel('average number of informed agents');
legend('pinfo=0.6')

%Représentation de l'évolution du nombre moyen de malades de degré k
%avec barres d'erreur
figure
plot(1:tmax,avnumbmalk)

```

```

title('time evolution of the average number of ill agents by degree')
xlabel('time');
ylabel('average number of ill agents');

%Representation de l'evolution de la proportion moyenne de malades
%de degre k avec barres d'erreur
figure
plot(1:tmax,avpropbmalk)
title('time evolution of the average proportion of ill agents by degree')
xlabel('time');
ylabel('average proportion of ill agents');

```

Enfin, voici les codes des deux fonctions de Timoteo Carletti utilisées lors de la détermination de la connexité d'un noeud.

```

function [Dia,Lm] = DiaLmL2(L,W)
% function [Dia,Lm] = DiaLmL2(L,W)
% This file computes some relevant functions on a given directed graph
% (possibly weighted)
%
% The Diameter (use W=L for lenghts measured in number of links, otherwise
% W_{ij} is the weight of the link ij), Dia : the maximal finite distance
% between any pair of nodes (cfr. Albert & Barabasi "Statistical Mechanics
% of Complex Networks", REVIEWS
% OF MODERN PHYSICS, VOLUME 74, JANUARY 2002)
% The mean path (use W=L for lenghts measured in number of links, otherwise
% W_{ij} is
% the weight of the link ij, Lm : average distance between any pair of
% nodes (cfr. Albert & Barabasi "Statistical Mechanics of Complex Networks",
% REVIEWS OF MODERN PHYSICS, VOLUME 74, JANUARY 2002)
%
% Inputs:
% L: adjacency list of the oriented graph (it should be of cell type)
% W: weight list (it should be of cell type), namely W{ii} contains the
% weights of the nodes belonging to L{ii}.
% If W is not passed then all these weigths are set equal to 1 and the
% graph is unweigthed
%
% Output:
% Dia : diameter
% Lm : mean path
%

```

```

% Required functions:
%
% dijkstraL2.m
%
% Timoteo Carletti, 8th may 2007
%

if ~iscell(L)
    fprintf(1,'warning L must be of cell type\n\r');
    fprintf(1,'nothing will be computed\n\r');
    Dia=0;
    Lm=0;
    return
end

n=length(L); % number of nodes

if (nargin==1)
    W=cell(size(L));
    for jj=1:n,
        W{jj}=ones(size(L{jj}));
    end
end

Dia=0; % initialize diameter
Lm=0; % initialize mean path

Dis = 0;
for s=1:n,
    % compute the total sum of (weighted) distances
    [d,S] = dijkstraL2(s,L,W);
    Dis = Dis+sum(d);
    % compute the diameter
    Dia = max(Dia,max(d));
end
% compute the mean path
Lm = Dis/(n*(n-1));

end

%-----

```

```

function [d,S] = dijkstraL2(s,L,W)
% function [d,S] = dijkstraL2(s,L,W)
% This is an implementation of the dijkstra algorithm, which detemines the
% shortest (or minimal cost) path between a given node s and all the
% remaining nodes.
% We assume positive weights and the graph given by its adjacency list.
%
% Inputs:
% s: source node index (integer between 1 and n (number of nodes)
% L: adjacency list of the oriented graph (it should be of cell type)
% W: weight list (it should be of cell type), namely W{ii} contains the
%   weights of the nodes belonging to L{ii}.
%   If W is not passed then all these weigths are set equal to 1 and the
% graph is unweigthed
%
% Output:
% d : distances from s : 0 means the node itself, inf means nodes not
% connected to s
% S : a possible path from s to any connectable node
%
% Required function
% none
%
% Timoteo Carletti, 7th may 2009
%

if ~iscell(L)
    fprintf(1,'warning L must be of cell type\n\r');
    fprintf(1,'distances will not be computed\n\r');
    S = [];
    d = [];
    return
end

n=length(L); % number of nodes

if (nargin==2)
    W=cell(size(L));
    for jj=1:n,
        W{jj}=ones(size(L{jj}));
    end
end

%d=ones(1,n)*inf; % initialize minimal dist function

```

```

d=ones(1,n)*10^10; % initialize minimal dist function. This should be Inf
% but I have to set it very large but finite to be able to perform the
% multiplication Inf*0=0
% (see line 73 : d(L{nodou})=actdist.*test+d(L{nodou}).*(1-test);)

S=[]; % initialize set S
Q=[1:n]; % initialize set Q

d(s) = 0;
Lq = length(Q);

% check that s is connected to at least one node
if isempty(L{s})
% no connection at all
    S=[s,S];
else
% at least one connection
    while (Lq > 0),
[du,u]=min(d(Q));
% remove the node u from the set Q
        nodou=Q(u);
        iu=find(Q==nodou);
        Im=[1:iu-1];
        Ip=[iu+1:Lq];
        Q=[Q(Im),Q(Ip)];
    %
    d(nodou)=du;
% add the node u to the set S
    S=[nodou,S];

% relax the distances of nodes v connected to u
        actdist=d(nodou) +W{nodou};
        test=(actdist < d(L{nodou}));
        d(L{nodou})=actdist.*test+d(L{nodou}).*(1-test);
% Note
% lines 69-71 perform faster the following task 74-81
% for v=1:n,
%         isvin=find(v==L{nodou}); % possible position of v in L{nodou}
% if ~isempty(isvin)
% if (d(nodou) +W{nodou}(isvin) < d(v))
%     d(v) = d(nodou)+W{nodou}(isvin);
% end
% end
% end
Lq = length(Q);
    end
end

```